



Contributions à la segmentation d'image : phase locale et modèles statistiques

Djamal Boukerroui

► To cite this version:

Djamal Boukerroui. Contributions à la segmentation d'image : phase locale et modèles statistiques. Traitement des images [eess.IV]. UTC Compiègne, 2014. tel-01170975

HAL Id: tel-01170975

<https://hal.science/tel-01170975>

Submitted on 2 Jul 2015

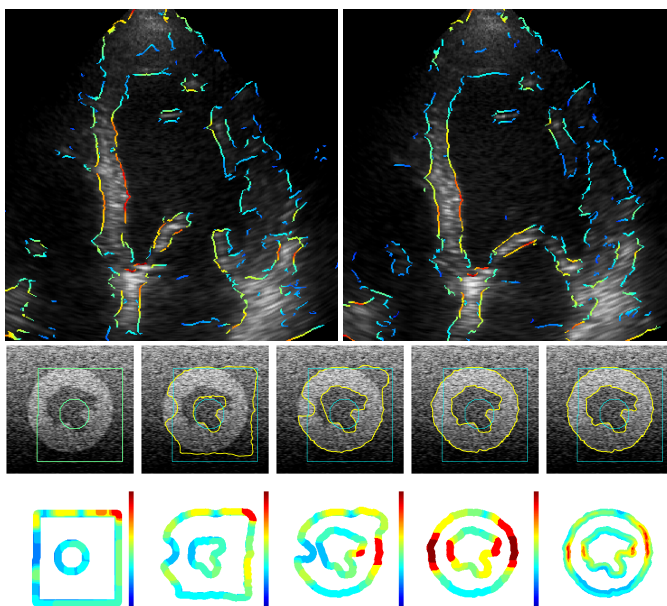
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Par Djamal BOUKERROUI

Contributions à la segmentation d'image :

phase locale et modèles statistiques



Mémoire présenté

en vue de l'obtention du diplôme
d'Habilitation à Diriger des Recherches

Contributions à la segmentation d'images information de phase et modèles statistiques

par

Djamal Boukerroui

Soutenue publiquement le 23 mai 2014 devant le jury composé de :

Jean-Louis COATRIEUX	Directeur de Recherche Émérite INSERM Laboratoire LTSI, Université de Rennes 1	Rapporteur
Christine FERNANDEZ-MALOIGNE	Professeur des Universités Laboratoire XLIM-SIC, Université de Poitiers	Rapporteur
Yuemin ZHU	Directeur de Recherche CNRS Laboratoire CREATIS, INSA-LYON	Rapporteur
Jean-Pierre COCQUEREZ	Professeur des Universités Émérite Laboratoire Heudiasyc, UTC	Examineur
Fabrice MERIAUDEAU	Professeur des Universités Laboratoire LE2I, Université de Bourgogne	Examineur

À ma femme Izumi
À ma fille Mina

“La pensée ne doit jamais se soumettre, ni à un dogme, ni à un parti, ni à une passion, ni à un intérêt, ni à une idée préconçue, ni à quoi que ce soit, si ce n’est aux faits eux-mêmes, parce que, pour elle, se soumettre, ce serait cesser d’être.”

Henri Poincaré

Remerciements

En tout premier lieu, je tiens à remercier Christine Fernandez-Maloigne, Jean-louis Coatrieux et Yuemin Zhu de m'avoir fait le grand honneur d'accepter de rapporter sur cette Habilitation à Diriger des Recherches. Je tiens également à remercier Fabrice Meriaudeau et Jean-Pierre Cocquerez pour leur participation à ce jury. Qu'ils trouvent tous ici l'expression de mon profond respect.

Une pensée à Bernard Dubuisson, qui n'a pas pu être des nôtres ce 23 mai 2014. Je le remercie très chaleureusement pour l'intérêt qu'il a toujours manifesté à mon travail.

Je remercie de nouveau Jean-Pierre pour ses encouragements permanents et surtout pour sa gentillesse et la très grande sympathie qu'il a toujours manifesté à mon égard.

D'autre part, je souhaite remercier tout particulièrement Sir Michael Brady et Alison Noble de l'Université d'Oxford, pour la confiance qu'ils m'ont accordée, pour leurs conseils avisés et l'enthousiasme dont ils ont toujours fait preuve. J'associe à ces remerciements Miguel Mulet Parada et Jérôme Declerck, mes partenaires de travail pendant mon séjour à Oxford.

Une partie des travaux présentés dans ce mémoire n'auraient pas vu le jour sans le travail des doctorants que j'ai eu l'honneur d'encadrer. Je les remercie pour la confiance qu'ils m'ont témoignée.

Je n'omettrai pas de remercier Vincent Frémont, Yves Grandvallet et Stéphane Mottelet, mes relecteurs. Merci pour votre aide si précieuse.

Enfin, je réserve mes remerciements les plus chaleureux à ma femme Izumi, mon relecteur d'anglais, pour son aide et son soutien permanent.



TABLE DES MATIÈRES

I	Rapport d'activité	11
	Curriculum vitæ	13
	État civil	13
	Poste actuel	13
	Cursus	13
	Parcours professionnel	14
	Activité d'enseignement	14
	Activités d'encadrement	14
	Activités de recherche	15
	Responsabilités administratives et fonctions d'intérêt collectif	16
	Synthèse des publications	17
1	Activités d'enseignement	19
1.1	Enseignements à UTC	20
1.2	Enseignements à Oxford	21
1.3	Enseignements à Lyon	21
2	Activités de recherche	23
2.1	Historique & résumé	23
2.2	Mobilité, collaborateurs internationaux	25
2.3	Projets nationaux et internationaux	25
3	Publications	27
	Revue avec comité de lecture	27
	Chapitres de livres	28
	Actes de congrès avec comité de lecture	28
	Rapports de recherche	29

Brevet	30
Mémoire de thèse	30
II Synthèse scientifique	31
1 Introduction	33
1.1 La segmentation d'image	33
1.2 Les contours actifs	36
1.2.1 Fonctionnelle de Mumford et Shah	37
1.2.2 Modèles locaux	38
1.3 Plan du document	39
2 Filtres passe-bande en quadrature	41
2.1 Introduction et contexte	41
2.1.1 Estimation de la phase et de l'amplitude locale 1D	43
2.1.2 Le cas multidimensionnel	44
2.2 Mesures de détection	45
2.2.1 La congruence de phase	45
2.2.2 Feature Asymmetry	46
2.3 Contributions	47
2.3.1 Synthèse et analyse de l'existant	47
2.3.2 Nouvelles familles de filtres	48
3 Rehaussement de contraste en Échocardiographie	51
3.1 Introduction	51
3.2 Interpolation de surface	52
3.2.1 Résolution dans l'espace ondelettes	52
3.3 Détection de contours	53
3.4 La méthode de correction	54
3.5 Résultats et bilan	54
4 Estimation de mouvement en Échographie	57
4.1 Mesure de similarité	58
4.2 L'approche de Singh	59
4.3 Contributions	59
4.4 Résultats et bilan	60
5 Segmentation d'image	61
5.1 Segmentation automatique des reins	61
5.1.1 Contexte et Problématique	61
5.1.2 Notre approche	61

5.1.3	Résultats	63
5.1.4	Bilan	64
5.2	Approches variationnelles en segmentation d'images échographiques . .	65
5.2.1	Introduction	65
5.2.2	Approche par phase locale	66
5.2.3	Approches statistiques locales	68
6	Schémas numériques pour le GVF	75
6.1	Formulation du GVF	76
6.2	Contributions	77
7	Bilan et perspectives	79
	Références	83
III	Principales publications	101
[5]	On the choice of band-pass quadrature filters	103
[12]	α Scale spaces filters	135
[8]	Enhancement of contrast in suboptimal ultrasound images	141
[25]	Velocity estimation in Ultrasound images	155
[10]	Identification des régions d'intérêts sur des images tomographiques	169
[3]	Phase based level set segmentation of ultrasound images	187
[16]	Ultrasound image segmentation using local statistics	199
[15]	A local Rayleigh model with spatial scale selections	205
[2]	Efficient numerical schemes for gradient vector flow	219

Première partie

Rapport d'activité



CURRICULUM VITÆ

État civil

Djamal Boukerroui
27 juillet 1972 à Bejaïa (Algérie)
Nationalités : Française, Algérienne
Vie maritale, pacsé, 1 enfant.

Poste actuel

Maître de conférences classe normale (CNU 61)
Laboratoire Heudiasyc, UMR CNRS 7253, Domaine Décision et Image
Université de Technologie de Compiègne
Tél : 03 44 23 52 79
Fax : 03 44 23 44 77
mel : djamal.boukerroui@hds.utc.fr
<http://www.hds.utc.fr/~dboukerr>

Cursus

- 1996–2000 **Thèse de Doctorat** en Signal Image Parole, (CREATIS) Institut National des Sciences Appliquées de Lyon.
 "Segmentation bayésienne d'images par une approche markovienne multirésolution. Application à l'imagerie échocardiographique et échographique 2D et 3D du sein".
 soutenue devant : M. Olivier Basset (directeur), M. Atilla Baskurt (co-directeur), M. Jean-Marc Chassery, Mme Alison Noble, M. Frédéric Truchetet (rapporteurs), M. Patrick Bouthemy (président), Mme. Isabelle Magnin (examineur).
- 1995–1996 **DEA** en Signal Image Parole, (CREATIS) Institut National des Sciences Appliquées de Lyon.
- 1990–1995 **Ingénieur** en Électronique, École Nationale Polytechnique d'Alger (Algérie).

1987–1990 **Baccalauréat**, Mathématique, Lycée Hafsa, Akbou (Algérie).

Parcours professionnel

depuis 2002 **Maître de conférences** au Département de Génie Informatique de l'Université de Technologie de Compiègne. Laboratoire Heudiasyc UMR CNRS 7253.

2000–2002 **Assistant de recherche** (Post-Doc), Medical Vision Laboratory MVL, Department of Engineering Science, Université d'Oxford (Angleterre).

1996 – 1999 **Vacataire** à l'IUT B, Université Claude Bernard - Lyon I.

1999 – 2000 **Vacataire** à l'ISTIL, Université Claude Bernard - Lyon I.

Activité d'enseignement (détails page 19)

depuis 2002 **Université de Technologie de Compiègne**

- SIT56, Niveau Master 2, (10h CM + 6h TD en moyenne de 2010 à 2013)
- TI11, Niveau Master 2, (10h CM en moyenne de 2002 à 2008)
- SY23, Niveau Branche, GI, (32h CM et 64h TD/an en 2002-04, 32h TD/an 2005-2008)
- SY06, Niveau Branche, GI, (32h TD/an, 2002 à 2005).
- SY02, Niveau Branche, GI, (32h à 192h TD selon les années, 2004-200x)
- SY16, Niveau Branche, GM, (8h TD, 2012).
- SY16, Niveau Branche, GM, (4h CM/an, 32h TD/an, 2006 à 2008).
- PS12, Niveau Branche, GB, (2h CM/an, 2008 à 2010).

2001-2002 **Oxford Brookes University**

- Vision par ordinateur, Niveau Licence, CM+TD+TP, (63h éq. TD).

1999-2000 **Université Claude Bernard- Lyon I, ISTIL**

- Robotique et vision, Niveau Licence, (32h TP).
- Traitement d'images, Niveau Licence, (16h TP).
- Traitement du signal, Niveau Licence, (16h TD).

1996-1999 **IUT B - Lyon I, Département Génie Industriel et Maintenance**

- Électronique, Niveau DUT, (TP, de 1996 à 1999, 144h éq. TD).
- Automatisme, Niveau DUT, (TP, de 1997 à 1999, 48h éq. TD).

Activités d'encadrement

Titulaire de la PEDR de 2006 à 2010 ;

J'ai suivi la formation **Encadrement du projet doctoral** en 2013.

Thèses

- 2009–2013 Q. YANG, *Segmentation d'images ultrasonores basée sur des statistiques locales avec une sélection d'échelles*, [18,16].
Taux d'encadrement 100%.
Financement : Bourse CSC.
Soutenance mars 2013.
Devenir : **post-doc**, Chine.
- 2008–2011 A. BELAID, *Segmentation des images ultrasonores par une approche de phase locale : Application aux images échocardiographiques pédiatriques*, [20,3,17].
Taux d'encadrement 80%, Co-direction avec J-F. Lerallut.
Financement : Bourse Région.
Soutenance le 28 novembre 2011.
Devenir : **Maître Assistant**, Université de Béjaïa (Algérie).
- 2005–2009 X. WAN, *Segmentation d'images : étude comparative de différentes approches de contours déformables*, [19].
Taux d'encadrement 50%, Co-direction avec J-P. Cocquerez.
Financement : Bourse CSC.
Soutenance le 19 octobre 2009.
Devenir : **Retour en Chine**.
- 2002–2006 W. TOUHAMI, *Identification et classification automatique de régions d'intérêt dans des images tomographiques : Application aux kystes du rein*, [23, 24,33, 22, 10].
Taux d'encadrement 80%, Co-direction avec J-P. Cocquerez.
Financement : Bourse État-Région.
Soutenance le 27 novembre 2006.
Devenir : **Maître Assistant**, Institut Supérieur des Sciences Appliquées et de Technologie de Mateur (Tunisie).

Master

- 2008 M. RODRIGUEZ SALCEDO *Réduction de bruit dans des images médicales*.
- 2005 R. BENMOKHTAR, *Segmentation par contour actif du type GVF : application à la segmentation des artères carotides en IRM*.
- 2004 H. OMRANI, *Analyse d'images mammographiques pour l'aide au diagnostic du cancer du sein : détection des microcalcifications*.

Activités de recherche

Thèmes de recherche : Traitement d'image médicale (segmentation, restauration, ultrasons)

- Étude théorique sur le choix de filtres passe-bande en quadrature pour leur utilisation dans des applications en détection.
- Segmentation et analyse d'images échocardiographiques : approches variationnelles et approches basées sur la phase locale.
- Détection automatique d'organe d'intérêt : application à la détection automatique des reins sur les images tomographiques.

Points forts :

- Traitement de données ultrasonore : expérience reconnue au niveau international.
- Très peu d'équipes travaillent sur les approches utilisant les filtres passe-bande en quadrature en France. Approches 'Local Energy' et 'Phase congruency' basées sur la généralisation de la transformée de Hilbert.

Projets nationaux et internationaux (détails page 25)

- 2012-2015 **Projet Phc Tassili : Scicibov**, *Système conjoint de compression et d'indexation basé-objet pour la vidéo*. Partenaire : UTC, INSA de Rennes, l'Université Djillali Liabès de Sidi Bel Abbès et l'Université de Mascara. Publication relative au projet : [14].
- 2011-2014 **Contrat de recherche** financé par le CETIM¹ (avec Vincent Frémont) sur la *Perception multi-capteurs pour la détection de risques de collision*. Le montant attribué est de **208 keuros**. Publication relative au projet : [13].
- 2008-2011 **Projet région Picardie : Échopédia**, *Développement d'outils informatiques en échocardiographie pédiatrique 3D pour l'exploration des cardiopathies congénitales*. Partenaires : UTC et CHU d'Amiens. Le montant de l'aide est de **26.258 €**. Publications relatives au projet : [20, 17, 3].
- 2002-2005 **Projet régional HTSC²** Système de communication pour la formation en médecine basé sur l'indexation d'images. Partenaires principaux : UTC, laboratoires Heudiasyc (thème diagnostic) et BioMécanique et BioIngénierie (thème imagerie), CHU d'Amiens, CHU de Lille et les sociétés Archimed et Planet HC. Le montant de l'aide : **448.919 €**. Publications relatives au projet : [23, 24, 33, 22, 10].
- 2000-2002 **Projet Européen Adequate (IST : 1999-10837)** *Assessment and Diagnosis in Echocardiography using Quantitative Techniques*, financé dans le cadre du 5^{ème} PCRD. Partenaires : Université d'Oxford, Hôpital Gregorio Marañón de Madrid, Le montant du projet était de **1.045.468 €**. Publications relatives au projet : [26, 5, 39, 8, 25].

Responsabilités administratives et fonctions d'intérêt collectif _____**Activité de relecture**

Membre de comités de révision d'articles pour des conférences internationales :

- depuis 2006 MICCAI : Medical Image Computing and Computer Assisted Intervention.
- depuis 2009 ISBI : IEEE International Symposium on Biomedical Imaging.
- 2003-2009 IPMI : Information Processing in Medical Imaging.
- 2006 CVAMIA : International workshop on Computer Vision Approaches to Medical Image Analysis.

1. Centre Technique des Industries Mécaniques.

2. Homme Technologie Systèmes Complexes.

- 2004 CVPR : Computer Vision and Pattern Recognition - Medical Imaging Section
- 2009 - 2010 IMAGAPP : International Conference on Imaging Theory and Applications
- depuis 2013 VISAPP : International Conference on Computer Vision Theory and Applications.

Révision d'articles pour publication dans des journaux internationaux. Par ordre de sollicitations :

- IEEE Transaction on Medical Imaging ;
- IEEE Transaction on Image Processing ;
- Medical Image Analysis ;
- IEEE Transactions on Biomedical Engineering ;
- IEEE Transactions on Information Technology in BioMedicine ;
- Pattern Recognition ;
- Traitement du signal ;
- Academic Radiology ;
- Signal, Image and Video Processing.

Travaux d'expertise

- Expertise de projet PEPS, campagnes 2009-2010, 2010-2011.
- Expertise de projet ANR, campagne 2010-2011.

Jury de Thèse

Examineur Benjamin Odry, 26 Oct. 2007, UTC ;

Rapporteur Ling Shao, 22 Juin 2007, Université d'Oxford ;

Basavaraj Amarapur, 28 Jan. 2011, Visvesvaraya Technological University, Inde (Rapport défavorable, pas de soutenance).

Responsabilités administratives

- Depuis 23-02-2012 : Membre élu (suppléant) du **conseil du laboratoire** Heudiasyc.
- Du 11-2011 au 11-2013 : Membre élu du **Bureau de département** du Génie Informatique de l'UTC.
- Membre élu de la **commission de spécialistes** 61^{ème} section, à l'UTC de 2005 à 2008.

Synthèse des publications

La liste complète de mes publications est donnée page 27. Elle se compose de

- 10 revues avec comité de lecture, dont 9 internationales ;
- 23 actes de congrès avec comité de lecture, dont 19 internationaux ;
- 1 chapitre de livre dans une édition internationale ;
- 4 rapports de recherche ;
- 1 brevet international.

ACTIVITÉS D'ENSEIGNEMENT

Je ne peux que me réjouir que, depuis toujours, les enseignements qu'il m'est arrivé de dispenser durant toute ma carrière, dans les quatre établissements, IUT B-Lyon I, ISTIL-Lyon et Brooks University (Oxford) en tant que vacataire, puis à l'UTC au sein du département du génie informatique, sont en parfaite adéquation avec mes compétences. Les publics visés vont de la première année post-bac à la dernière année de formation d'ingénieur, mais aussi au niveau M2 du master TIS¹ de l'UTC. Les différents enseignements auxquels j'ai participé sont détaillés dans les paragraphes suivants. La table 1.1 résume les enseignements effectués à l'UTC.

TABLE 1.1 – Récapitulatif des enseignements à l'UTC en heures équivalent TD. R=responsabilité de l'UV, C=cours magistral, D=travaux dirigés, T=travaux pratiques.

	SY23	SY06	SY16	SY02	TI11	SIT56	PS12
12/13	-	-	D - 8h	D - 192h	-	CD - 19h	-
11/12	-	-	-	D - 192h	-	CD - 19h	-
10/11	-	-	-	D - 160h	-	CD - 16h	C - 3h
09/10	RCD - 80h	-	-	D - 92h	-	-	C - 3h
08/09	RCD - 80h	-	CD - 40h	D - 64h	-	-	C - 3h
07/08	RCD - 85h	-	D - 34h	D - 64h	RC - 15h	-	-
06/07	RCD - 85h	T - 10h	D - 34h	D - 138	RC - 15h	-	-
05/06	RCD - 119h	T - 16h	-	D - 68h	RC - 15h	-	-
04/05	RCD - 119h	D - 34h	-	D - 34h	C - 18h	-	-
03/04	RCD - 119h	D - 34h	-	-	C - 12h	-	-
02/03	RCD - 119h	D - 34h	-	-	C - 9h	-	-

1. Technologies de l'Information et des Systèmes

1.1 Enseignements à UTC

Depuis mon recrutement en tant que Maître de conférences à l'UTC en 2002, je suis intervenu dans les Unités de Valeur (UV) suivantes :

- 2002-2010 Vision & Image** - (code UTC : SY23)- **Responsable d'UV**. Cours et Travaux dirigés. Formation initiale ingénieur de l'UTC (niveau branche). L'objectif de cette UV est de donner aux étudiants une introduction aux outils et méthodes classiques de traitement d'images. Les principaux thèmes abordés sont : la couleur, la transformée de Fourier, échantillonnage, quantification, filtrage, amélioration d'images, détection de points de contour, segmentation région, introduction à la morphologie mathématique et à l'analyse de texture. Le cours s'appuie sur de nombreux travaux pratiques afin de familiariser l'étudiant avec les problèmes spécifiques de l'image numérique. Les TDs se faisaient initialement avec le logiciel Khoros sous Unix. J'ai dû faire une première migration sous Windows en utilisant le logiciel VisuQuest en 2006, puis une seconde adaptation vers Matlab® en 2010 s'appuyant principalement sur le livre de [Gonzalez et al. \(2009\)](#).
- 2002-2008 Traitement et analyse d'images**- (TI11) - **responsable de l'UV de 2005 à 2008**. Niveau Master 2. Après un rappel des notions fondamentales sur les processus aléatoires et de quelques traitements de bas niveau en image, le cours aborde essentiellement le problème de la segmentation et celui de la restauration d'image et cela dans deux formalismes différents (variationnel puis markovien). La partie variationnelle est introduite par Jean-Pierre Coccoquerez, et je m'occupe de la partie markovienne.
- 2010-2013 Vision et traitements embarqués** - (SIT56). Niveau Master 2. L'objectif de ce cours est d'offrir aux étudiants des connaissances de base en vision par ordinateur pour des traitements embarqués tels que le filtrage d'images, la segmentation, la reconstruction 3D par stéréovision, l'estimation du mouvement, et la reconnaissance d'objets, pour des tâches telles que l'odométrie visuelle, la détection d'obstacles et l'aide à la conduite. Cette UV est une refonte de TI11 pour l'orienter davantage vers la vision 3D et son application à la robotique. Vincent FRÉMONT s'occupe de la partie "vision 3D", et moi de la partie "image". Après un rappel des notions fondamentales de filtrage et de la transformée de Fourier (sous toutes ses formes), le cours aborde essentiellement les problèmes de la restauration d'image, de la détection de point de contour, et de l'estimation de mouvement. En plus des cours et des travaux dirigés, les étudiants travaillent par binôme sur des mini-projets qui leur permettent d'appliquer les notions vues en cours et de coder certains algorithmes sous Matlab.
- 2002-2007 Analyse et traitement du signal** - (SY06) -TD & TP, formation initiale ingénieur de l'UTC (niveau branche). L'objectif de cette UV est de donner aux étudiants les éléments de théorie utilisés pour construire les outils d'analyse des signaux (complément sur l'analyse de fonctions, décomposition sur des bases orthogonales, modèles aléatoires) et pour comprendre les traitements élémentaires basés en particulier sur les représentation de Fourier (filtrage, transformée de Fourier, estimation spectrale).
- 2006-2008 Les applications du traitement du signal** - (SY16) -TD, formation initiale ingénieur de l'UTC (niveau branche). L'UV a pour objectif d'offrir à de fu-

turs ingénieurs les moyens d'appliquer les méthodes avancées du traitement du signal. L'accent est mis sur l'art du traitement, de l'analyse et de l'interprétation des signaux. L'approche se veut pragmatique et est illustrée par de nombreuses applications sur des signaux réels empruntés à divers domaines (parole, musique, vibroacoustique, biomédical...). Aussi, les étudiants ont la possibilité de choisir eux mêmes le sujet du mini-projet sur lequel ils vont travailler.

depuis 2004 **Méthodes Statistiques pour l'ingénieur** - (SY02) -TD, formation initiale ingénieur de l'UTC (niveau TC & branche). L'UV introduit les concepts et les méthodes de base de la statistique en vue de son utilisation dans les sciences de l'ingénieur (estimation, tests d'hypothèses, analyse de la variance, régression linéaire, intervalle de confiance).

depuis 2002 Suivi des étudiants pendant leur stage assistant ingénieur (TN09) et leur projet industriel de fin d'études (TN10); Durée : 6 mois. volume 40h éq. TD par an. Le TN09 se déroule en milieu professionnel, du secteur public ou du secteur privé. Située au 3^{ème} semestre de branche (les élèves ingénieurs sont bac+3). Le TN10 est réalisé en milieu professionnel au cours du dernier semestre du cursus ingénieur (soit Bac+ 4,5). Il met l'étudiant en condition d'exercice de ses futures fonctions dans le milieu professionnel qu'il rejoindra. Une soutenance orale et un rapport écrit sont les éléments de validation pour les deux stages.

1.2 Enseignements à Oxford

En 2001-2002, j'ai eu la responsabilité de l'enseignement du module **Vision par Ordinateur** à des élèves ingénieurs à Oxford Brookes University, School of Computing and Mathematical Sciences. Les neuf séances de cours de 2 heures abordaient les points élémentaires suivants :

- Acquisition d'images : types de projection (perspective et orthographique);
- Image binaire et traitement : opérations morphologiques, étiquetage et classification d'objet binaire;
- Convolution et détection de contours;
- Vision stéréo et flot optique : notions de base, problèmes de correspondance, détection de coins;
- Reconnaissance de forme : notions de base, étiquetage de lignes, étude du système de vision SCERPO développé par David G. Lowe.

1.3 Enseignements à Lyon

Dans l'année scolaire 1999-2000, j'étais **vacataire à l'Université Claude Bernard- Lyon I, ISTIL**. Les enseignements dispensés sont les suivants :

- **Robotique et vision** (TP, niveau : 3^{ème} année, 32 heures). Au cours de ces travaux pratiques, j'ai appris aux étudiants à utiliser deux systèmes industriels de vision par ordinateur (CIVIM et le système robotique ADEPT1). Le but de cet enseignement était de montrer des exemples d'applications industrielles de la vision par ordinateur, en insistant en particulier sur les contraintes de temps

réel et de robustesse aux conditions d'acquisitions des images (illumination, champ de vue).

- **Traitement d'images** (TP, niveau : 3^{ème} année ; 16 heures). Encadrement d'une vingtaine d'étudiants au cours de 4 séances de 4 heures de travaux dirigés de traitement d'image avec programmation en langage C. Le travail demandé aux étudiants était la réalisation de quelques algorithmes de base de traitement d'images : visualisation d'une image (gris/couleur), transformée de Fourier 2D, filtrage fréquentiel, produit de convolution, filtrage d'ordre, détection de contour, opérations morphologiques. . .
- **Traitement du signal** (TD, niveau 2^{ème} année, 16 heures). Introduction au traitement du signal. Les sujets traités étaient : classification des signaux (énergie/puissance finie), transformée de Fourier, fonctions d'autocorrélation et d'intercorrélation, transformée de Hilbert et notion de signal analytique.

Durant ma thèse, j'ai été **vacataire à IUT-B Lyon I** au Département Génie Industriel et Maintenance. Les enseignements dispensés sont les suivants :

1996-1999 Electronique (TP ; niveau : 1^{ère} année ; 144 h éq. TD au total). Les travaux pratiques permettaient d'aborder les points suivants : appareils de mesure, circuits passifs, diodes, transistors bipolaires et montages associés, amplificateurs opérationnels en régime linéaire et non-linéaire, oscillateurs, transistors à effet de champ, alimentations stabilisées et réalisation de circuits imprimés.

1997-1999 Automatismes (TP ; niveau : 1^{ère} année ; 48 h éq TD au total). Après un cours d'introduction aux différentes notions d'automatisme nécessaires aux travaux pratiques (2 heures), les 6 séances de travaux pratiques de 4h portaient sur la logique séquentielle câblée, les automatismes pneumatiques et la programmation d'automates industriels sous PL7-2 (avec le langage Grafcet).

ACTIVITÉS DE RECHERCHE

2.1 Historique & résumé

Accueilli en octobre 1995 au sein du laboratoire CREATIS (UMR CNRS 5515) de l'INSA de Lyon dans le cadre de mon stage de DEA, j'ai tout d'abord travaillé, sous la direction du professeur Robert UNTERREINER, sur le thème de la fusion de signaux physiologiques pour l'estimation de la qualité de préparation du geste sportif [34]. En thèse sous la direction de MM. Olivier BASSET et Atila BASKURT, je me suis intéressé au problème de la segmentation d'images échographiques 2D et 3D. J'ai développé une approche "région" formulée dans un cadre statistique bayésien avec un *a priori* markovien [30, 29, 28, 27, 9, 6, 7].

Dans le cadre de mon post-doc – de février 2000 à août 2002 – à l'université d'Oxford, j'ai travaillé sur la segmentation et l'analyse de séquences d'images échocardiographiques. Mes axes de recherche portaient sur la segmentation d'images par congruence de phase, le rehaussement de contraste [26, 8, 38], l'estimation de mouvement [25, 37] et l'étude des propriétés des filtres en quadrature [36, 5, 11] dédiés à des tâches de détection ou d'estimation de mouvement. Mes recherches avaient pour objectif de faire collaborer ces différentes techniques de traitement d'images au problème de la segmentation de données fortement bruitées. J'ai également collaboré avec le Docteur Timor KADIR, chercheur au laboratoire MVL d'Oxford, sur un problème de sélection de paramètre saillant pour la description d'image [35]. Mes recherches étaient financées en partie par le projet européen IST-FP5 ADEQUATE.

En septembre 2002, je suis membre du laboratoire Heudiasyc de l'UTC et j'ai intégré le thème ASTRID, devenu le domaine DI (Décision & Image), afin de renforcer l'équipe Image, plus précisément imagerie médicale. De 2002 à 2006, j'ai travaillé en collaboration avec Jean-Pierre COCQUEREZ, sur la problématique de l'indexation d'images médicales. D'un point de vue méthodologique, mon activité dans le cadre de ce projet concerne la partie image, plus précisément le problème de la segmentation et de l'identification automatique des régions d'intérêts, utiles pour le calcul de la signature de l'index. Notre activité est soutenue par le projet "Système de communication pour la formation en médecine basé sur l'indexation d'images" dans le cadre de l'axe mobilisateur du contrat de plan état-région "Homme, Technologie, Système Complexe" (HTSC). J'ai dirigé la thèse de Mlle W. TOUHAMI, ma première, avec J.P COCQUEREZ sur cette thématique de fouille

d'images par le contenu avec une application aux images tomographiques de reins. Notre principale contribution concerne l'approche originale que nous avons proposée pour l'identification automatique des reins. Nos travaux sur ce sujet ont fait l'objet de plusieurs publications [33, 23, 24, 22, 10].

A partir de 2008, j'ai recentré ma recherche sur le traitement d'images échographiques, ma thématique de recherche avant l'intégration d'Heudiasyc. Ainsi, j'ai pu monter avec Jean-François LERALLUT et le Docteur Yves MAINGOURD au CHU d'Amiens le projet région ECHOPEDIA. La thèse de M. A. BELAID co-dirigée avec JF. LERALLUT, financée dans le cadre de ce projet, aborde la problématique de la segmentation d'images ultrasonores par une approche de phase locale. L'originalité de notre approche était l'association d'un détecteur de contours basé sur la phase avec un algorithme de contours actifs dans le formalisme des ensembles de niveaux. Ce travail s'appuie sur mes travaux antérieurs sur les filtres en quadrature [5, 11] et est valorisé dans [20, 3, 17]. La thèse de Mlle Q. YANG s'inscrit également dans cette thématique. Arrivée en octobre 2009 dans le cadre du programme CSC¹, elle a fait sa thèse sous ma direction, toujours sur la problématique de segmentation d'images échographiques. Ici nous avons opté pour une approche variationnelle statistique. La problématique de recherche porte sur l'utilisation des statistiques locales ; plus précisément sur la sélection de la taille (l'échelle) du noyau qui définit la localité spatiale. Une solution originale est proposée dans [18, 16].

Enfin, suite à l'évaluation du mon laboratoire en 2011, l'activité "Imagerie Médicale" ne figure plus dans le projet du laboratoire Heudiasyc pour 2012-2015². J'ai été donc encouragé à rejoindre les activités "Vision". Ce qui a modifié les champs d'applications de mes travaux. Cela m'a conduit à m'impliquer dans deux projets en cours avec des applications Vision (détails de la liste des projet en page 25). Je participe notamment à l'encadrement de la thèse de Manh Tuan BUI sur l' '*Étude d'un système de perception multi-capteurs pour la détection de risques de collision*'.

Mes recherches, qui s'inscrivent dans le domaine du traitement d'image, ont toujours eu pour objet la **résolution d'un problème de segmentation**, avec des applications en imagerie médicale et plus particulièrement les images ultrasonores. Les spécificités de ces données ont orientées fortement les approches utilisées, sans toutefois affaiblir l'aspect générique de mes contributions. Sommairement, les contributions de mes travaux d'après thèse portent sur :

- l'amélioration d'images [26, 8] ;
- la prise en compte d'information *a priori* [33, 24, 23, 22, 19, 10] ;
- les modèles d'attaches aux données de type frontière [11, 20, 1, 12, 5, 3] et statistique [18, 16, 31, 17, 15] ;
- la segmentation de séquence d'images par le mouvement [25] ;
- la mise en œuvre de schémas numériques [32, 21, 2] utiles en segmentation ;
- des travaux de synthèse [5, 4, 11] et autre [35, 14, 13].

Une synthèse scientifique des principaux travaux est donnée dans la seconde partie de ce document.

1. China Science Council.

2. Manque de visibilité, très faible taille de l'équipe.

2.2 Mobilité, collaborateurs internationaux

En 2000–2002, j’ai effectué un stage post-doctoral, sous la direction de Alison Noble, au Medical Vision Laboratory de l’université d’Oxford, dirigé par Michael Brady. Durant mon séjour, j’ai collaboré avec mon ex-collègue de bureau au CREATIS de l’INSA de Lyon, Marc Robini, sur la problématique d’interpolation de surface. Cette dernière est une étape importante de la méthode de rehaussement de contraste que j’ai développé [8].

Une fois revenu en France et ayant intégré le laboratoire Heudiasyc, j’ai collaboré avec Alison Noble, dans un premier temps pour finir des travaux entamés [25, 11], puis pour écrire une synthèse sur la segmentation d’images ultrasonores [4].

Aujourd’hui, je collabore avec mon ex. doctorant Ahror Belaid, enseignant-chercheur à l’université de Béjaïa sur la problématique des filtres en quadrature et leurs applications en segmentation d’image [12].

2.3 Projets nationaux et internationaux

Les différents projets auxquels j’ai participé sont détaillés dans cette section par ordre chronologique, du plus récent au plus anciens.

2012-2015 Projet Phc Tassili : Sccibov, *Système conjoint de compression et d’indexation basé-objet pour la vidéo*. Partenaire : UTC, INSA de Rennes, l’Université Djillali Liabès de Sidi Bel Abbès et l’Université de Mascara. Le premier objectif du projet consiste à étudier le codage vidéo scalable basé région. Le second objectif concerne l’indexation dans le domaine compressé. Le projet étant très orienté compression, mon implication se limite à la participation à l’encadrement de la thèse de M. Ahmed Boudaib (Université de Mascara) sur l’extraction de fond sur des séquences d’images. Cette problématique est l’une des sous tâches du premier objectif.
Publications relatives au projet : [14].

2011-2014 Contrat de recherche financé par le CETIM³ (avec Vincent Frémont). Ce contrat de recherche s’inscrit dans une problématique de *perception multi-capteurs pour la détection de risques de collision*. L’idée est ici de se concentrer sur la perception visuelle embarquée et de développer une méthodologie d’analyse d’images multi-capteurs et de fusion de données 3D multimodales pour la reconnaissance de situations dangereuses et de risques de collision entre un engin de chantier et des personnes. Ce contrat englobe le financement d’une thèse (étudiant Manh Tuan Bui sous la direction de V. Frémont) et un contrat d’accompagnement. Dans le cadre de ce contrat, je participe à l’encadrement de la thèse de M. Bui. J’interviens sur les aspects image notamment l’extraction de caractéristique et des points saillants.
Le montant attribué est de **208 keuros**. Publications relatives au projet : [13]

2008-2011 Projet région Picardie : Echopedia, *Développement d’outils informatiques en échocardiographie pédiatrique 3D pour l’exploration des cardiopathies congénitales*. Partenaires : UTC et CHU d’Amiens. L’objectif est le développement de méthodes de restauration et de segmentation d’images d’échographie pédiatrique tridimensionnelles pour l’exploration des cardiopathies congénitales.

3. Centre Technique des Industries Mécaniques.

Chez l'adulte, les cardiopathies sont des anomalies de la vascularisation myocardique, alors que l'architecture est normale. Chez le nouveau né, la physiologie cardiaque est modifiée par des anomalies de ses structures anatomiques. Dans le cadre de ce projet, j'ai été le responsable scientifique des deux tâches du projet et le directeur (à 80%) de la thèse de M. A. Belaid.

Le montant de l'aide est de **26.258 €**.

Publications relatives au projet : [20, 17, 3].

2002-2005 **Projet régional HTSC⁴** Système de communication pour la formation en médecine basé sur l'indexation d'images. Partenaires principaux : UTC, laboratoires HEUDIASYC (thème diagnostic) et Biomécanique et GBM (thème imagerie), CHU d'Amiens, CHU de Lille et les sociétés Archimed et Planet HC.

Le projet repose sur l'idée qu'une requête basée sur une image est plus informative et performante qu'une requête purement textuelle et évite l'étape fastidieuse de description de l'image par des mots clé. L'objectif est de concevoir un système de communication pour la formation initiale et permanente en médecine basé sur l'indexation d'images. Le produit final est le démonstrateur d'un système de formation interactif reliant des bases de données d'images médicales, des utilisateurs (étudiants ou médecins) à des fins de formation et d'aide au diagnostic. C'est dans le cadre de ce projet que j'ai eu l'occasion d'encadrer ma première thèse, celle de W. Touhami. La problématique recherche est l'identification automatique des reins (sains et pathologiques) sur des images tomographiques bidimensionnelles (CT). Les deux difficultés majeures sont d'une part la contrainte du tout automatique des traitements imposée par l'application de recherche d'image par le contenu et d'autre part la nature des données.

Le montant de l'aide : **448.919 €**.

Publications relatives au projet : [23, 24, 33, 22, 10].

2000-2002 **Projet Européen Adequate (IST : 1999-10837)** *Assessment and Diagnosis in Echocardiography using Quantitative Techniques*, financé dans le cadre du 5^{ème} PCRD. Partenaires : Université d'Oxford, Hôpital Gregorio Marañon de Madrid, l'Hôpital John Radcliffe d'Oxford et la société Mirada Solution^{Ltd}. L'objectif du projet est de concevoir un prototype pour le diagnostic automatique des données échocardiographiques. Cela implique le développement de nouveaux algorithmes pour l'extraction de caractéristiques et de suivi d'objets sur des images échographiques, et leur intégration dans un système de diagnostic convivial pour le clinicien en vue d'une utilisation clinique pour l'aide au diagnostic de maladies du cœur.

Dans le cadre de ce projet, j'ai été post-doc sous la direction de Alison Noble à l'Université d'Oxford. Mes recherches portaient sur le développement d'outils de segmentation automatique de données échocardiographiques. J'étais responsable des livrables de 4 workpackages. J'ai notamment participé aux évaluations du projet par la commission européenne à Bruxelles.

Le budget du projet était de **1.045.468 €**.

Publications relatives au projet : [26, 5, 39, 8, 25].

4. Homme Technologie Systèmes Complexes.

PUBLICATIONS

Citations Google Scholar : <http://scholar.google.com/citations?hl=fr&user=aJvMT38AAAAJ>

TABLE 3.1 – Résumé des publications les plus citées selon Google Scholar au 15/01/2014.

Référence	[4]	[6]	[25]	[5]	[3]
Nombre de citations	476	95	77	64	46

Revues avec comité de lecture

Revues internationales

- [1] A. Belaid and **D. Boukerroui**, "A new generalised α scale spaces quadrature filters," *Pattern Recognit.*, à paraître, 2014.
- [2] **D. Boukerroui**, "Efficient numerical schemes for gradient vector flow," *Pattern Recognit.*, vol. 45, no. 1, pp. 626–636, 2012.
- [3] A. Belaid, **D. Boukerroui**, Y. Maingourd, and J.-F. Lerallut, "Phase based level set segmentation of ultrasound images," *IEEE Trans. Info. Tech. Biomed.*, vol. 15, no. 1, pp. 138–147, 2011.
- [4] J. A. Noble and **D. Boukerroui**, "Ultrasound image segmentation : A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, 2006.
- [5] **D. Boukerroui**, J. A. Noble, and M. Brady, "On the choice of band-pass quadrature filters," *J. Math. Imag. Vision*, vol. 21, no. 1, pp. 53–80, 2004.
- [6] **D. Boukerroui**, A. Baskurt, J. Noble, and O. Basset, "Segmentation of ultrasound images—multiresolution 2D and 3D algorithm based on global and local statistics," *Pattern Recognit. Lett.*, vol. 24, no. 4-5, pp. 779–790, 2003.
- [7] **D. Boukerroui**, O. Basset, A. Baskurt, and G. Gimenez, "A multiparametric and multiresolution segmentation algorithm of 3-D ultrasonic data," *IEEE Trans. Ultra. Fer. Freq. Control*, vol. 48, pp. 64–77, Jan. 2001.
- [8] **D. Boukerroui**, J. A. Noble, M. C. Robini, and M. Brady, "Enhancement of contrast regions in suboptimal ultrasound images with application to echocardiography," *Ultrasound Med. Biol.*, vol. 27, pp. 1583–1594, Dec. 2001.

- [9] **D. Boukerroui**, O. Basset, N. Guérin, and A. Baskurt, "Multiresolution texture based adaptive clustering algorithm for breast lesion segmentation," *European Journal of Ultrasound*, vol. 8, no. 2, pp. 135–144, 1998.

Revues nationales

- [10] **D. Boukerroui**, W. Touhami, and J. Cocquerez, "Identification automatique des régions d'intérêts sur des images tomographiques rénales," *Traitement du Signal*, vol. 26, pp. 239–254, mai 2009.

Chapitres de livres

- [11] **D. Boukerroui**, J. A. Noble, and M. Brady, *Frontiers in Robotics Research*, ch. On the Selection of Band-Pass Quadrature Filters, pp. 67–111. Nova Publishers, Max A. Denket ed., 2006.

Actes de congrès avec comité de lecture

Congrès internationaux

- [12] A. Belaid and **D. Boukerroui**, " α scale spaces filters for phase based edge detection in ultrasound images," in *IEEE Int. Symp. Biome. Imaging*, Beijing, China, Apr. 2014, p. à paraître.
- [13] M.-T. Bui, V. Fremont, **D. Boukerroui**, and P. Letort, "People detection in heavy machines applications," in *Int. Conf. on Cybernetics and Intelligent System & Robotics, Automation and Mechatronics (CIS-RAM)*, (Manila, Philippines), p. to appear, Nov 2013.
- [14] K. Kpalma, C. Bai, M. Chikr El-Mezouar, K. Belloulata, N. Taleb, L. Belhallouche, and **D. Boukerroui**, "A new histogram-based descriptor for images retrieval from databases," in *Advances in Intelligent Analysis of Medical Data and Decision Support Systems* (J. Kacprzyk, ed.), vol. 473 of *Studies in Computational Intelligence*, (Sofia, Bulgarie), pp. 101–112, Jul 2013.
- [15] **D. Boukerroui**, "A local Rayleigh model with spatial scale selection for ultrasound image segmentation," in *British Machine Vision Conference*, (Surrey, Royaume-Uni), pp. 84–84, Sep 2012.
- [16] Q. Yang and **D. Boukerroui**, "Ultrasound image segmentation using local statistics with an adaptive scale selection," in *IEEE Int. Symp. Biome. Imaging*, (Spain), pp. 1096–1099, IEEE, May 2012.
- [17] A. Belaid, **D. Boukerroui**, Y. Maingourd, and J.-F. Lerallut, "Implicit active contours for ultrasound images segmentation driven by phase information and local maximum likelihood," in *IEEE Int. Symp. Biome. Imaging*, (Chicago Illinois, USA), pp. 630–635, IEEE, Apr. 2011.
- [18] Q. Yang and **D. Boukerroui**, "Optimal spatial adaptation for local region-based active contours : An intersection of confidence intervals approach," in *IMAGAPP* (G. Csurka, M. Kraus, and J. Braz, eds.), (Algarve, Portugal), pp. 87–93, SciTePress, March 5-7 2011.
- [19] X. Wan, **D. Boukerroui**, and J.-P. Cocquerez, "Image segmentation based on kernel PCA and shape prior," in *3rd Int. Conf. on Digital Image Processing*, Chengdu, Chine, Apr 2011.
- [20] A. Belaid, **D. Boukerroui**, J.-F. Lerallut, and Y. Maingourd, "Phase based level set segmentation of ultrasound images," in *IEEE Int. Conf. on Information Technology and Applications in Biomedicine (ITAB'2009)*, Larnaca, Cyprus, 2009.
- [21] **D. Boukerroui**, "Efficient numerical schemes for gradient vector flow," in *IEEE Int. Conf. Image Proc.*, (Cairo, Egypte), pp. 4057–4060, Nov. 2009.
- [22] **D. Boukerroui**, W. Touhami, and J.-P. Cocquerez, "Automatic regions of interest identification and classification in CT images : Application to kidney cysts," in *IEEE First Workshops on Image Processing Theory, Tools and Applications*, pp. 257–264, 2008.
- [23] W. Touhami, **D. Boukerroui**, and J. P. Cocquerez, "A statistical approach for automatic kidneys detection," in *IEEE Int. Conf. Image Proc.*, pp. 740–743, 2005.

- [24] W. Touhami, **D. Boukerroui**, and J.-P. Cocquerez, "Fully automatic kidneys detection in 2D CT images : A statistical approach," in *MICCAI* (J. S. Duncan and G. Gerig, eds.), vol. 3749 of *Lecture Notes in Computer Science*, pp. 262–269, Springer, 2005.
- [25] **D. Boukerroui**, J. A. Noble, and M. Brady, "Velocity estimation in ultrasound images : A block matching approach," in *Info. Proc. Med. Imaging* (C. J. Taylor and J. A. Noble, eds.), vol. 2732 of *Lecture Notes in Computer Science*, pp. 586–598, Springer, 2003.
- [26] **D. Boukerroui**, J. A. Noble, and M. Brady, "Feature enhancement in low quality images with application to echocardiography," in *Info. Proc. Med. Imaging* (M. F. Insana and R. M. Leahy, eds.), vol. 2082 of *Lecture Notes in Computer Science*, pp. 453–460, Springer, 2001.
- [27] **D. Boukerroui**, O. Basset, A. Baskurt, and J. A. Noble, "Segmentation of echocardiographic data. Multiresolution 2D and 3D algorithm based on gray level statistics," in *MICCAI* (C. J. T. 0002 and A. C. F. Colchester, eds.), *Lecture Notes in Computer Science*, (Cambridge, England), pp. 516–524, Springer-Verlag, 1999.
- [28] **D. Boukerroui**, O. Basset, and A. Baskurt, "Multiresolution adaptive image segmentation based on global and local statistics," in *IEEE Int. Conf. Image Proc.*, pp. 358–361, 1999.
- [29] **D. Boukerroui**, O. Basset, A. Baskurt, J.-M. Gorce, D. Friboulet, and G. Gimenez, "A segmentation algorithm of 3D ultrasonic data based on tissue characterization," in *IEEE Proceedings of Ultrasonics Symposium*, vol. 2, pp. 1349–1352 vol.2, 1998.
- [30] **D. Boukerroui**, O. Basset, A. Baskurt, A. Hernandez, N. Guerin, and G. Gimenez, "Texture based adaptive clustering algorithm for 3D breast lesion segmentation," in *IEEE Proceedings of Ultrasonics Symposium*, vol. 2, pp. 1389–1392 vol.2, 1997.

Congrès francophones

- [31] **D. Boukerroui**, "Segmentation d'images échographiques par la méthode des ensembles de niveaux avec un modèle de Rayleigh local," in *Traitement et Analyse de l'Information : Méthodes et Applications*, (Hammamet, Tunisie), pp. 231–238, Oct 2011.
- [32] **D. Boukerroui**, "Schémas numériques de résolution du flot de vecteur gradient," in *Traitement et Analyse de l'Information : Méthodes et Applications*, (Hammamet, Tunisie), pp. 523–532, May 2009.
- [33] W. Touhami, **D. Boukerroui**, and J. Cocquerez, "Détection automatique de reins sur des images tomographiques 2D," in *Traitement et Analyse de l'Information : méthodes et applications (TAIMA)*, (Tunisie), pp. 149–154, 2005.
- [34] **D. Boukerroui**, H. Rada, E. Vernet-Maury, and R. Unterreiner, "Fusion de signaux physiologiques par contrôle flou pour l'estimation de la qualité de la préparation du geste sportif," in *Rencontres Francophones sur la Logique Floue et ses Applications*, Cépaduès Editions, pp. 357–363, 1997.

Rapports de recherche

- [35] T. Kadir, **D. Boukerroui**, and M. Brady, "An analysis of the scale saliency algorithm," Technical Report OUEL 2264/03, Oxford University, 2003.
- [36] **D. Boukerroui** and J. A. noble, "On band-pass quadrature filters", *Technical Report. MVL*, Oxford University, 50p, Mars 2002.
- [37] **D. Boukerroui** and J. A. noble, "Optic flow estimation", *Adequate progress report 2*, (CONFIDENTIAL), Jun. 2001, 7p.
- [38] **D. Boukerroui** and J. A. noble, "Feature enhancement", *Adequate progress report 1*, (CONFIDENTIAL), Nov. 2000, 7p.

Brevet

- [39] **Dr. D. Boukerroui** and Dr. J. A. Noble, "Improvements in image velocity estimation", *Brevet No 0228300.0, British Patent Application*, 4 décembre 2002, Extension au USA W02004052016 2004-06-17.

Mémoire de thèse

- [40] **D. Boukerroui**, "Segmentation bayésienne d'images par une approche markovienne multi-résolution. Application à l'imagerie échocardiographique et échographique 2D et 3D du sein," Thèse de Doctorat, Institut National Des Sciences Appliquées de Lyon, 2000.

Deuxième partie

Synthèse scientifique

INTRODUCTION

Dans ce qui suit, après une définition de la problématique de la segmentation d'image, je présenterai dans un premier lieu la vision moderne variationnelle de la segmentation d'image et ses liens à d'autres problématiques de traitement d'image, plus particulièrement, la restauration, le recalage d'images et l'estimation du mouvement. Dans un second lieu, je présenterai une synthèse historique des travaux en lien avec cette vision moderne de la segmentation d'image. Cela donnera un éclairage sur mes contributions, après thèse, dans ce domaine.

1.1 La segmentation d'image

Quand l'Homme interprète visuellement une scène, il fait appel à des connaissances de haut niveau. Il a une capacité extraordinaire à faire le lien entre ce qu'il observe et ce qu'il a déjà vu auparavant, et à extraire de scènes très complexes les objets qui l'intéressent. Il est maintenant établi que les deux stratégies de traitement, ascendante (*bottom-up*) et descendante (*top-down*) sont nécessaires et utilisées dans ce processus d'interprétation. Néanmoins, l'interprétation automatique des images par ordinateur est un problème d'une extrême complexité nécessitant plusieurs étapes. Dans ce contexte, on fait souvent la distinction entre les *traitements de bas niveau* et les *traitements de haut niveau*. Les premiers travaillent directement sur les valeurs attachées aux pixels de l'image, sans forcément faire le lien avec la réalité qu'elles représentent. Tandis que les seconds opèrent sur des entités symboliques constituant une interprétation de la description de l'image élaborée par les traitements de bas niveaux.

L'interprétation visuelle d'une scène exige la capacité à reconnaître les objets et leur localisation dans l'image. Le premier objectif représente le problème de reconnaissance de forme et le second celui de la segmentation d'image. Dans sa première définition, la *segmentation* d'images peut être considérée comme un traitement de bas niveau. Elle consiste en une partition du domaine spatial Ω de l'image \mathcal{I} en M régions $\{\Omega_1, \Omega_2, \dots, \Omega_M\}$ connexes, disjointes et **homogènes**. Autrement dit, elle tente de représenter l'image sous la forme de régions disjointes afin de faciliter l'interprétation et aider à la compréhension de la scène imagée. La segmentation n'est donc pas une fin en soi, mais une étape importante, et souvent nécessaire dans la chaîne de traitement. Elle conditionne le succès ou l'efficacité de tout système qui a pour ambition de donner une représentation synthétique du contenu de l'image ou une interprétation de la scène.

La richesse des travaux de recherche réalisés sur ce sujet souligne d’une part l’importance de cette tâche en traitement d’image et d’autre part sa complexité et la difficulté à la résoudre. En effet, la segmentation d’image est un problème mal posé, au sens de Hadamard ([Hadamard 1932](#)), difficile à formaliser, à résoudre et à mettre en œuvre. La diversité et la complexité des images augmentent fortement la difficulté du problème. Il est d’autant plus difficile lorsque les données à traiter sont de nature complexe telles par exemple les images médicales acquises dans des conditions cliniques. La littérature sur le sujet est très abondante. De nombreuses synthèses méthodologiques ([Angenent et al. 2006](#), [Arbelaez et al. 2011](#), [Bresson et Chan 2008](#), [Cremers et al. 2007](#), [Freixenet et al. 2002](#), [Papari et Petkov 2011](#), [Peng et al. 2013](#), [Zhang et Lu 2001](#), [Ziou et Tabbone 1998](#)), spécifiques aux données médicales ([Angelini et al. 2005](#), [Gordillo et al. 2013](#), [Heimann et Meinzer 2009](#), [Katouzian et al. 2012](#), [McInerney et Terzopoulos 1996](#), [Noble et Boukerroui 2006](#), [Olabarriaga et Smeulders 2001](#), [Tarachandani et Boltz 2006](#), [Tavakoli et Amini 2013](#), [Werghi 2007](#)) et livres ([Aubert et Kornprobst 2006](#), [Blake et Isard 1998](#), [Cocquerez et Philipp-Foliguet 1995](#), [Klette et Zamperoni 1996](#), [Mitiche et Ben Ayed 2010](#), [Morel et Solimini 1995](#), [Osher et Paragios 2003](#), [Paragios et al. 2005](#), [Szeliski 2010](#)) traitent de ce sujet.

La segmentation peut être fondée directement sur le niveau de gris ou la couleur de chaque pixel, ou bien sur un ou plusieurs attributs estimés dans le voisinage du pixel. Dans la définition donnée précédemment, le terme clés est le mot “homogène”. Ainsi, classiquement les méthodes de segmentation sont principalement classées en deux approches. Une *approche région* qui regroupe les pixels de l’image ayant des propriétés communes, qui recherche ou caractérise donc l’homogénéité des régions. Une seconde *approche*, dite *contour*, duale à la première, qui recherche plutôt les transitions entre les régions, autrement dit la rupture de l’homogénéité. Ces deux approches ont dominé la littérature sur le sujet pendant des décennies.

Les principales difficultés des approches classiques, incluent notamment la séparation de l’objet d’intérêt en plusieurs régions et la fusion des parties de l’objet avec le fond. Ces limitations sont dues à des ambiguïtés inévitables puisque la plupart des objets ne sont pas homogènes en termes de couleur, texture, etc. En outre, les objets d’intérêts dans l’image ne sont pas nécessairement contrastés avec le fond, entraînant ainsi de potentielles fusions entre eux. Autrement dit, d’une façon plus formelle, le modèle de description des données observées est incapable de lever ces ambiguïtés afin de fournir la partition d’intérêt. Ce problème ne peut être résolu sans une connaissance *a priori* sur la classe d’objets considérée. Ainsi, des approches *top-down* sont apparues. Une classe importante de cette catégorie d’approches inclut les modèles déformables, détaillés ci-dessous. Plus récemment, des solutions aspirant à résoudre le problème conjoint de détection et de segmentation d’objets procèdent généralement par combinaison des deux stratégies, ascendante et descendante, sont proposées ([Borenstein et Ullman 2008](#), [Fidler et al. 2013](#), [Levin et Weiss 2009](#), [Yang et al. 2012](#)).

Par ailleurs, [Morel et Solimini \(1995\)](#) montrent dans une synthèse sur la segmentation d’image, que malgré la grande diversité des approches existantes de segmentation, elles peuvent être formalisées comme un problème variationnel. En outre, ils montrent que la grande majorité d’entre-elles tendent à minimiser explicitement ou non la fonctionnelle de Mumford et Shah, présentée pour résoudre conjointement les problèmes de la segmentation et de la restauration d’image ([Mumford et Shah 1985, 1989](#)). Quelques années plus tard, [Mumford \(1994\)](#) a donné une interprétation bayésienne de la fonctionnelle. La segmentation d’image variationnelle est formalisée dans [Morel et Solimini \(1988\)](#). [Koepfler, Lopez, et Morel \(1994\)](#) ont proposé un algorithme de segmentation

rapide multi-échelles du type “fusion de région” minimisant une approximation simplifiée de la fonctionnelle de Mumford et Shah. Cependant l'étude mathématique de la fonctionnelle s'avéra une tâche difficile, et est à notre connaissance toujours d'actualité. Plusieurs approximations sont proposées dans la littérature et résumées dans (Aubert et Kornprobst 2006). Le lecteur peut consulter (Pock et al. 2009) pour un exemple de contribution récente sur le sujet.

En imagerie médicale, le problème de segmentation est de plus en plus lié à un autre problème de traitement, aussi difficile à résoudre que la segmentation : le recalage d'images (Angelini et al. 2005). Cela est apparu lors de l'introduction des modèles de segmentation intégrant des connaissances *a priori* via des modèles de forme et d'apparence (Heimann et Meinzer 2009, Lu et al. 2011, et références citées) ou sous forme d'atlas (Cabezas et al. 2011). Nous citerons ces quelques contributions d'application d'atlas en segmentation (Gorthi et al. 2011, Iglesias et al. 2013, Lorenzo-Valdes et al. 2004, Park et al. 2003, Rohlfing et al. 2004, Vemuri et al. 2003). Ainsi le problème de segmentation est résolu via la résolution du problème de recalage ou les deux problèmes sont traités conjointement. Le terme hybride “*Regmentation*”, combinaison du terme anglais Registration (recalage) et Segmentation a été proposé récemment pour désigner cette classe de méthodes (Erdt et al. 2012).

Enfin, le problème de segmentation de séquence d'images est intrinsèquement lié au problème d'estimation de mouvement ou de suivi d'objets. Nous distinguons donc deux problèmes, celui de la segmentation du champs de vitesse¹ et celui du suivi d'objets. Ainsi, le problème de segmentation est vu principalement comme un problème de suivi (eg. Boukerroui et al. 2003b, Comaniciu et al. 2004, Freedman et Zhang 2004, Mansouri 2002, Zhou et al. 2005) ou un problème de recalage spatio-temporel (voir eg. Ledesma-Carbayo et al. 2005, Rueckert et al. 2002) ou une formulation conjointe suivi-segmentation (eg. Dydenko et al. 2006, Mignotte et al. 2001, Mikić et al. 1998, Paragios 2003, Paragios et Deriche 2000, Zhou et al. 2005), mouvement-segmentation (Cremers 2006, Fazekas et al. 2009, Mitiche et Sekkati 2006, Paragios et Deriche 2005, Vazquez et al. 2006) est adoptée. Le lecteur peut aussi consulter un certain nombre de livres où des chapitres entiers sont consacrés à cette problématique (Aubert et Kornprobst 2006, Mitiche et Ben Ayed 2010, Osher et Paragios 2003, Paragios et al. 2005).

On peut souligner que la formulation du problème de segmentation dans un cadre statistique offre généralement une meilleure prise en compte du processus de formation d'image. Aussi, comme il a déjà été précédemment relevé par Morel et Solimini (1995), la plupart des méthodes de segmentation peuvent être formulées comme un problème de minimisation d'une fonction d'énergie. De plus une interprétation bayésienne existe généralement. Les algorithmes de segmentation diffèrent souvent dans leur mise en œuvre et dans les choix des contraintes de régularité des solutions recherchées, imposées par des *a priori* génériques ou spécifiques à un problème de segmentation donné ; et cela d'une façon supervisée ou non. Aussi, le formalisme variationnel a été largement utilisé ces deux dernières décennies car il permet de définir un cadre unifié pour résoudre la plupart des problèmes de traitement d'image et de définir explicitement la fonction d'énergie à optimiser. Des formulations alternatives du problème de segmentation existent bien sûr (voir eg. Arbelaez et al. 2011, Szeliski 2010).

1. À différencier des méthodes qui utilisent la segmentation d'image comme contrainte pour améliorer l'estimation du champs de vitesse (voir eg. Black et Jepson 1996, Xu et al. 2008).

1.2 Les contours actifs

L'introduction des contours déformables en segmentation d'image par [Kass, Witkin, et Terzopoulos \(1988\)](#) est certainement une rupture par rapport aux approches classiques de segmentation d'image qui proposent des solutions par le biais d'un processus ascendant et séquentiel. L'objectif est de détecter les contours d'un objet d'intérêt dans une image et non pas de définir une partition de l'image en différentes zones homogènes. L'objet d'intérêt n'est pas nécessairement une zone "homogène" au sens classique. L'idée de base est la formulation du problème de segmentation comme un problème de minimisation d'une énergie qui permet de faire évoluer une courbe vers les caractéristiques désirées, telles les frontières de l'objet d'intérêt. C'est un mécanisme très puissant qui permet l'introduction d'une façon élégante et générique des informations *a priori* de haut niveau. Ces dernières sont combinées avec des caractéristiques image bas-niveau. Ainsi, nous pouvons disposer d'informations de forme, de taille ou encore de la distribution d'intensité à l'intérieur ou à l'extérieur de l'objet.

Ce travail pionnier qui a permis l'introduction du modèle "snake", un *contour actif paramétrique*, est à l'origine du développement de toute une famille de méthodes de segmentation d'image dénommée, *contours actifs* ou modèles déformables. Cette famille a connu un intérêt particulier de la part de la communauté de l'imagerie médicale pour plusieurs raisons. Nous pourrions citer par exemple leur grande flexibilité, leur richesse à intégrer des connaissances *a priori* ou la possibilité de faire de la segmentation interactive ([Liang et al. 2006](#), [McInerney et Terzopoulos 1996](#), [Olabarriaga et Smeulders 2001](#)). Néanmoins, la formulation originale des contours actifs souffre de plusieurs inconvénients. En effet, dans les premiers travaux, l'énergie image est une fonction du gradient de l'image, donc très locale et la mise en œuvre utilise une paramétrisation explicite de la courbe. Cela a pour conséquence :

1. une sensibilité du modèle au bruit et à l'initialisation ;
2. la paramétrisation du contour pose des difficultés de mise en œuvre mais surtout, ne permet pas le changement de topologie du contour initial.

Les inconvénients dus à la paramétrisation ont été solutionnés par l'introduction des contours *géométriques* ([Caselles et al. 1993](#)) et leur mise en œuvre via les ensembles de niveaux (level sets) ([Malladi et al. 1995, 1996](#), [Osher et Sethian 1988](#)). Le premier point est essentiellement lié à la nature du terme d'attache aux données basé contour des premiers modèles de contours actifs. Plusieurs solutions sont proposées dans la littérature que l'on peut classer en deux catégories. Une première catégorie, regroupe les méthodes toujours basées frontière et qui ont pour objectif soit la recherche du minimum global de l'énergie (voir eg. [Cohen et Kimmel 1997](#)) ou d'augmenter la zone d'attractivité du contour actif ; nous citerons par exemple ([Cohen et Cohen 1993](#), [Cohen 1991](#), [Li et Acton 2007](#), [Paragios et al. 2004](#), [Xie 2010](#), [Xu et Prince 1998b](#)). Toutefois ces méthodes restent toujours sensibles au bruit et il a été constaté que les énergies basées frontière possèdent beaucoup de minima locaux (voir eg. [Cohen 1997](#), [Cremers et al. 2007](#)). Les limitations des approches frontières suggèrent naturellement d'orienter les recherches vers des solutions utilisant une information plus globale de l'image et intégrant la nature statistique du bruit. Les méthodes de la seconde catégorie permettent donc une meilleure prise en compte du modèle de bruit de l'image. Ce sont les méthodes de *contours actifs* dits "*région*" qui exploitent la richesse des modélisations statistiques variées déjà existantes, apparues suite aux travaux de [Geman et Geman \(1984\)](#). Les travaux pionniers dans ce sens sont dus à [Cohen et al. \(1993\)](#), [Ronfard \(1994\)](#), [Zhu et Yuille \(1996\)](#) et [Chakraborty](#)

et al. (1996) suivi par d'importantes contributions (Chan et Vese 2001, Chesnaud et al. 1999, Jehan-Besson et al. 2003, Paragios et Deriche 2002, Yezzi et al. 2002).

L'introduction de critères basés région dans les modèles déformables a certes permis d'augmenter leur robustesse par rapport au bruit mais elle a aussi permis de s'affranchir des problèmes d'initialisation, conduisant ainsi à des méthodes de segmentation automatiques. Les premiers travaux s'appuient généralement sur des modélisations statiques paramétriques *globales* de l'image, région par région (Chan et Vese 2001, Chesnaud et al. 1999, Jehan-Besson et al. 2003, Paragios et Deriche 2002, Samson et al. 2000). Cette classe de méthodes est la plus riche (Aujol et al. 2003, Ayed et al. 2006, 2005, Cremers et al. 2007, Feng et al. 2013, Heiler et Schnörr 2005, Lecellier et al. 2010, Martin et al. 2004b, Sarti et al. 2005, Sawatzky et al. 2013). Des critères non-paramétriques sont apparus ensuite (Adam et al. 2009, Aubert et al. 2003, Herbulot et al. 2006, Hibbard 2004, Houhou et al. 2009, Joshi et Brady 2010, Kadir et Brady 2003, Kim et al. 2005, Martin et al. 2006, Michailovich et al. 2007, Ni et al. 2009, Peyre et al. 2012, Zhang et al. 2012).

Cela dit, l'aspect global de la prise d'information des critères région pose des difficultés lorsque l'hypothèse d'homogénéité est violée par exemple par la présence d'un dégradé d'intensité. Les minima (globaux ou locaux) de l'énergie ne correspondent alors pas souvent à la solution désirée. Notons que les méthodes de type frontières, au delà des critiques formulées ci-dessus, présentent une meilleure robustesse dans ce cas précis vu qu'elles sont quasi-invariantes par rapport aux dégradés d'intensité.

1.2.1 Fonctionnelle de Mumford et Shah

Plusieurs modèles de contours actifs géométriques sont directement dérivés de la fonctionnelle de Mumford et Shah (Brox et Cremers 2007, 2009, Chan et al. 2006, Chan et Vese 2001, Chan et al. 2007, Law et al. 2008, Le et Vese 2007, Tsai et al. 2001, Vese et Chan 2002). Parmi ces méthodes, nous trouvons des modèles constants par morceaux (Bae et al. 2011, Bresson et al. 2007, Chan et al. 2006, Chan et Vese 2001, Le et Vese 2007, Vese et Chan 2002) ou lisses par morceaux (Bresson et al. 2007, Brox et Cremers 2007, 2009, Chen et al. 2013b, Le et Vese 2007, Tsai et al. 2001, Vese et Chan 2002) et des travaux sur la minimisation globale du modèle de contour actif obtenu (Bresson et al. 2007, Chan et al. 2006, Chen et al. 2013b, Law et al. 2008).

La recherche de l'optimum global a conduit les chercheurs à abandonner la représentation implicite de la carte de segmentation par les fonctions ensembles de niveaux et de se rapprocher des représentations classiques utilisées en classification multi-labels afin d'exploiter au mieux les récentes avancées en optimisation convexe (Bae et al. 2011, Chambolle et Pock 2011, Gu et al. 2012, Yildizoglu et al. 2012). Une solution alternative est proposée par Boykov et Kolmogorov (2003) où le problème de recherche de la courbe géodésique minimale sous une métrique riemannienne arbitraire est transformé en un problème de coupe sur un graphe. Ainsi le minimum global est obtenu par le biais d'un algorithme de minimisation globale de type "graph-cut" (Boykov et Funka-Lea 2006, Boykov et Jolly 2001, Kolmogorov et Zabih 2004). Cette idée est encore étendue dans (Boykov et al. 2006, Kolmogorov et Boykov 2005). Une approche similaire est proposée dans (Kirsanov et Gortler 2004). L'utilisation de la technique du graph-cut dans des algorithmes de segmentation par contours actifs s'est vulgarisée (El-Zehiry et Elmaghraby 2008, Tao 2012, Tao et al. 2010, Tao et Tai 2011, Xu et al. 2007, Zeng et al. 2006, 2008). Une comparaison récente des approches issues des modélisations continues, conduisant à la résolution d'Équations aux Dérivées Partielles (EDPs), et discrètes, issues de modélisation de type Markovienne, est apparue dans (Nieuwenhuis et al. 2013).

1.2.2 Modèles locaux

Le modèle de Mumford et Shah lisse par morceaux a suscité un intérêt particulier car il permet de tolérer des variations lisses à l'intérieur des régions. Cela offre donc plus de flexibilité en relaxant la contrainte d'*homogénéité* des régions et permet de solutionner efficacement les problèmes liés aux changements d'illumination de la scène en imagerie optique, à l'atténuation en imagerie ultrasonore ou à l'inhomogénéité du champ magnétique en Imagerie par Résonance Magnétique (IRM). Cependant, les premières solutions de ce modèle dans le cadre des contours actifs nécessitent la résolution d'un système d'EDPs pour obtenir les fonctions d'approximation (Bresson et al. 2007, Tsai et al. 2001, Vese et Chan 2002). Une solution efficace consiste à rechercher une représentation des fonctions d'approximation dans une base de fonctions, pas nécessairement lisses, comme proposé initialement par Vazquez et al. (2004, 2006). Cette modélisation est réapparue (peut-être indépendamment) dans (Mahmoodi et Sharif 2007) et plus récemment dans (Chen et al. 2013a). Une solution alternative est apparue par l'introduction récente des modèles statistiques paramétriques locaux (Boukerroui 2012, Brox et Cremers 2007, 2009, Darolti et al. 2008, El-Zehiry et Elmaghraby 2008, Faisal et Pluempitiwiriawej 2012, Lankton et Tannenbaum 2008, Li et al. 2011a, 2008, Liu et Zhang 2013, Piovano et al. 2007, Wang et al. 2009a, Wu et Yang 2012, Yang et Boukerroui 2011, Yang et al. 2010) et non paramétriques (Chen et Guo 2007, Guo et Chen 2006). Parmi ces travaux, Brox et Cremers (2009) ont montré que la minimisation de la fonctionnelle lisse par morceaux de Mumford-Shah est équivalente à une approximation de premier ordre de la Maximisation A Posteriori (MAP) d'un modèle bayésien supposant un bruit blanc gaussien local. Ici les fonctions d'approximation du modèle lisse par morceau sont données par l'estimation des moyennes locales du modèle de bruit.

Une seconde classe de modèles locaux utilise une modélisation dite par '*patch*'. L'idée est d'utiliser toute l'information locale autour d'un pixel donné comme vecteur d'observation, autrement dit tout le bloc de pixels centré autour du pixel courant. Cette représentation est utilisée initialement dans des algorithmes de synthèse de texture, puis elle est devenue très populaire suite à l'introduction de l'algorithme de débruitage, dit des moyennes non locales, de Buades et al. (2005a,b). En segmentation d'image, Bresson et Chan (2008) sont les premiers à proposer plusieurs algorithmes de segmentation non supervisée variationnelle s'appuyant sur les différences de patch. Werlberger et al. (2012) propose un algorithme efficace de minimisation d'une version multi-labels du modèle de Pott non-local proposé dans (Bresson et Chan 2008). Houhou et al. (2009) ont proposé d'appliquer cette idée pour estimer un descripteur de texture, et l'utiliser dans un algorithme de segmentation non paramétrique. Récemment, Jung et al. (2011, 2012) ont proposé une nouvelle classe de contours actifs '*non-local*' basés sur une distance entre bloc de pixels. Une application de ce modèle est proposée dans (Liu et al. 2013).

Enfin, d'autres propositions, dites de bande étroite, limitent la définition de l'énergie à minimiser sur une bande étroite autour du contour à l'instant t , dénommée région active (Li et Yezzi 2007). Les motivations de cette technique diffèrent selon les auteurs. Certaines contributions sont motivées par des contraintes de complexité et de temps d'exécution et d'autres cherchent la robustesse aux dégradés d'intensité. Aussi nous citerons des méthodes optimisées par les graph-cut (El-Zehiry et Elmaghraby 2008, Tao 2012, Tao et al. 2010, Tao et Tai 2011, Xu et al. 2007, Zheng et al. 2013) ou par des algorithmes issus d'une formalisation continue du problème (Li et Yezzi 2007, Mille 2009, Yildizoglu et al. 2012).

Ces travaux récents montrent un meilleur comportement des modèles locaux sur

des images présentant des dégradés d'intensité. Il est important de souligner que l'utilisation des *statistiques locales* (donc de modèles de type région) en segmentation d'image sont apparus auparavant dans un cadre bayésien classique ; à notre connaissance initialement par Pappas (1992), suivi par (Ashton et Parker 1995, Boukerroui et al. 2003a, 1999, 1998, Unser 1995). Aussi une lecture attentive du travail de Ronfard (1994), un des pionniers de l'intégration des statistiques région dans ce contexte, révèle que celles-ci sont estimées localement. À cet égard, le travail de Kadir et Brady (2003) peut aussi être interprété comme la première introduction des statistiques locales dans le formalisme des courbes de niveaux. En effet, le terme d'attache aux données à minimiser n'est que la somme des entropies croisées entre l'histogramme local et l'histogramme global de chaque région.

1.3 Plan du document

Le contexte historique ayant déjà été présenté, les sources de financement citées et mes principaux collaborateurs et laboratoires d'accueils rappelés, et après une courte introduction sur la problématique de la "segmentation d'image", fil conducteur de mes recherches, la synthèse de celles-ci est l'objet du reste du document. Notre choix est de faire une synthèse concise. Dans la mesure du possible, nous avons adopté les mêmes notations que les articles joints au document vers lesquels le lecteur pourra se référer pour plus de détails et de résultats.

Le chapitre 2 présente essentiellement les éléments théoriques nécessaires pour l'estimation des propriétés locales des signaux 1D (phase et énergie) et 2D (phase, énergie et orientation) en utilisant des filtres en quadrature. Un résumé de mes contributions dans ce domaine clos ce chapitre.

Le chapitre 3 présente une méthode d'amélioration de contraste sur des séquences d'images échocardiographiques. Elle repose essentiellement sur un détecteur de contour en utilisant l'information de phase.

Le chapitre 4 présente une méthode d'estimation de mouvement sur des images échographiques qui prend en compte la nature statistique du bruit sur ce type d'images. L'approche est développée dans un but de suivi de contour et par conséquent pour résoudre un problème de segmentation.

Le chapitre 5 présente pour une grande partie les recherches effectuées sous ma direction portant sur la segmentation d'image.

Le chapitre 6 présente mon travail sur les techniques numériques pour la résolution du problème de flot de vecteur gradient, une technique très utile dans beaucoup d'applications en traitement d'image et notamment la segmentation d'image.

Le dernier chapitre fournira quelques conclusions et perspectives de recherche.

Dans tout le document, mes travaux sont cités entre crochets par des numéros qui font référence à la liste de mes publications. Les principales (cités par numéros en gras) sont reproduites à la fin du document, dans l'ordre de citation dans la synthèse scientifique. Les autres références bibliographiques (sous forme auteur et année de publication) sont regroupées à la page 83.

FILTRES PASSE-BANDE EN QUADRATURE

2.1 Introduction et contexte

Le système visuel humain a toujours été une source d'inspiration pour le développement d'algorithmes de vision par ordinateur. Fait intéressant, plusieurs expériences physiologiques ont suggéré que des structures simples, comme les lignes, les bords et les jonctions, jouent un rôle important dans notre système de vision. Ainsi, ces caractéristiques, certes de bas niveau, ont toujours été considérées comme un élément central dans l'analyse par vision par ordinateur. Leur détection est donc une opération fondamentale qui doit être traitée de manière fiable et robuste. L'extraction de telles caractéristiques image a été largement étudiée dans la littérature et demeure encore un domaine de recherche actif (Bourennane et al. 2002, Canny 1986, Demarcq et al. 2011, Elder et Zucker 1998, Felsberg et Sommer 2004, Haralick 1984, Konishi et al. 2003, Kovesei 1999, 2003, Lindeberg 1998, 2008, Marr et Hildreth 1980, Mcilhagga 2011, Olson 2000, Owens et al. 1989, Papari et Petkov 2011, Pellegrino et al. 2004, Poggio et al. 1988, Shen et Castan 1992, Sumengen et Manjunath 2005, von Gioi et al. 2010, Wietzke et al. 2008, Ziou 1991, Ziou et Tabbone 1998).

L'étude de cette vaste littérature suggère que les outils de traitement les plus classiques sont basés sur l'amplitude du signal. Tout au long de l'histoire du traitement numérique de l'image, le lissage et le calcul de dérivées spatiales ont été des sujets d'intenses études (Ando 2000, Krueger et Phillips 1989). La grande diversité d'opérateurs optimaux proposés en témoigne. La détection des frontières, de type ligne ou échelon a reçu une attention particulière (Bourennane et al. 2002, Canny 1986, Cocquerez et Philipp-Foliguet 1995, Haralick 1984, Lindeberg 1998, Mcilhagga 2011, Papari et Petkov 2011, Pellegrino et al. 2004, Poggio et al. 1988, Shen et Castan 1992, Ziou 1991, Ziou et Tabbone 1998). Il est admis que la différentiation est sensible et amplifie le bruit. L'influence de ce dernier peut heureusement être réduite par la sélection d'une échelle appropriée pour la fonction de lissage. Les mesures basées sur l'amplitude restent cependant très dépendante du rapport signal à bruit et sont sensibles aux variations de contraste dans l'image. En outre, une image contient plusieurs types de frontières (et pas seulement de type échelon) et les caractéristiques intéressantes de l'image n'ont souvent pas la forme idéalisée supposée par les modèles. Cela peut donc réduire considérablement l'optimalité de la détection et peut en sus conduire à des réponses multiples

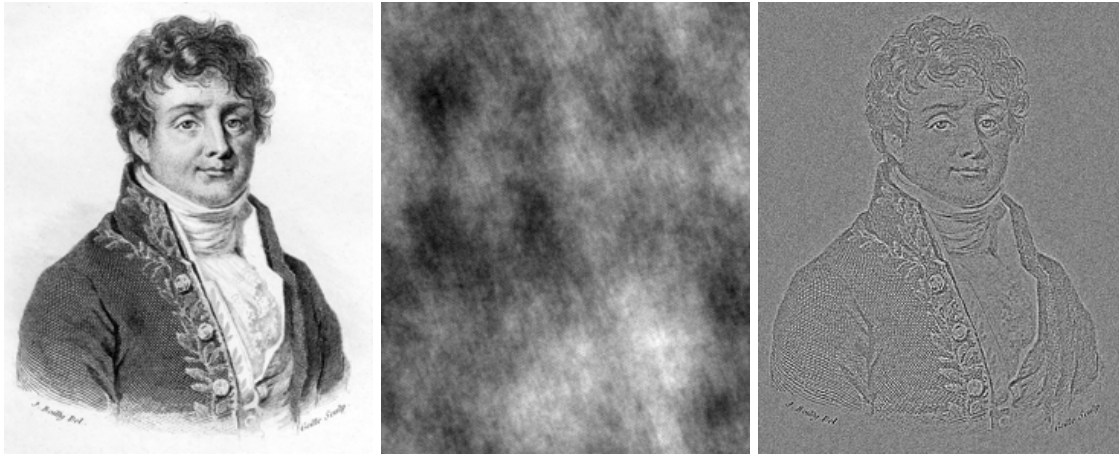


Figure 2.1 – Importance de la phase. De gauche à droite : Gravure de Julien Léopold Boilly, image de Joseph Fourier (image de wikipédia), reconstruction avec phase aléatoire, reconstruction avec amplitude constante.

imprévisibles. En conséquence, les contributions récentes ont atteint un haut degré de sophistication (Papari et Petkov 2011). Nous trouvons des méthodes automatiques de sélection de l'échelle d'analyse (Elder et Zucker 1998, Lindeberg 1998, Ng et Bharath 2004), et des approches basées sur les théories des espaces échelles linéaires (Lindeberg 1994, Witkin 1983) et non-linéaire (Alvarez et al. 1993, Perona et Malik 1990b) et des modèles alternatifs de détection, par exemple statistiques (Konishi et al. 2003).

Une alternative aux approches basées amplitude est l'utilisation de l'information de la phase. L'importance de cette dernière en traitement d'image est mise en évidence par Oppenheim et Lim (1981). Ils ont démontré que cette composante est cruciale pour la perception des structures sur des images. Ils ont établi qu'à partir d'une image donnée, si on en dérive deux images synthétiques, dont l'une contient la phase (de Fourier), et l'autre l'amplitude, il en ressort incontestablement que c'est l'image contenant la phase qui sera toujours perçue. La figure 2.1 illustre cette expérience.

La détection d'attributs en utilisant l'information de phase a été intensivement étudiée suite à la publication du modèle de détection basé sur l'énergie locale (Morrone et Owens 1987). Ce modèle postule que les discontinuités sont perçues aux points dans une image où les composantes de Fourier sont en phase. Il a également été montré que ces points correspondent à des maxima de l'énergie locale (Venkatesh et Owens 1990). De plus, le modèle nous renseigne sur le type de structures détectées, permettant ainsi leurs classifications. En effet, mesurer la phase locale ou encore la *congruence de phase* sur plusieurs échelles est un moyen de caractériser les différences d'intensités en terme de forme d'intensité et non pas en terme de force d'intensité. En effet, l'amplitude du signal (ou le gradient) nous informe sur la *force* de la discontinuité, alors que la phase locale nous informe sur sa *forme*, ce qui rend cette dernière théoriquement indépendante des changements d'intensité, donc robuste en pratique.

Cette observation a permis le développement d'un nombre d'algorithmes de détection basés sur la *phase locale* (Bharath et Ng 2005, Felsberg et Sommer 2000, 2004, Heitger 1995, Kovess 1996, 1999, 2003, Reisfeld 1996, Robbins et Owens 1997, Ronse 1993, 1997, Sedlazeck 2008, Venkatesh et Owens 1990). Ce formalisme possède un avantage certain puisque il offre un cadre unifié et général de détection de tout type de discontinuité (pas seulement de type échelon ou ligne). De nombreux avantages de l'utilisation de mesures

basées sur la phase ont été rapportés. Peut-être la propriété la plus souhaitable est son invariance (théorique) au contraste de l'image.

2.1.1 Estimation de la phase et de l'amplitude locale 1D

Une des méthodes les plus populaires pour estimer les propriétés locales des signaux réels mono-dimensionnels est basée sur la représentation analytique du signal. Il s'agit d'un signal complexe donné par l'équation suivante (Gabor 1946, Picinbono 1997) :

$$f_A(x) = f(x) + if_{\mathcal{H}}(x) \quad (2.1)$$

où $i = \sqrt{-1}$ et $f_{\mathcal{H}}(x)$ est la transformée de Hilbert du signal $f(x)$. Le passage dans le domaine de Fourier nous conduit à :

$$F_A(\omega) = F(\omega) \cdot [1 + \text{sign}(\omega)] . \quad (2.2)$$

Par conséquent, le signal analytique correspondant à f est obtenu en supprimant toutes ses fréquences négatives et en multipliant les fréquences positives par deux. Il convient de noter qu'aucune information sur f n'est perdue en raison de la symétrie hermitienne du spectre d'une fonction réelle. L'*amplitude locale* et la *phase locale* ou 'instantanée' sont définies d'une manière simple à partir de cette représentation complexe de $f(x)$ comme étant le module et l'argument du signal analytique :

$$f_A(x) = A(x)e^{i\varphi(x)} . \quad (2.3)$$

Cependant, un calcul direct de ces grandeurs locales ne nous sera d'aucune utilité en pratique. C'est principalement parce que la transformée de Hilbert/signal analytique est définie sur l'ensemble du support/spectre du signal, tandis que, par exemple, pour la détection de primitives, la localisation dans l'espace et en fréquence sont hautement souhaitables. En effet, si notre signal se compose de signaux partiels à des échelles différentes, une seule phase et amplitude locales ne peuvent représenter adéquatement le signal. Un outil classique permettant ce type de traitement est la transformée de Fourier à court terme (voir eg. Mallat 1999, chap.4). Une alternative est de procéder à l'estimation des propriétés locales d'un signal d'approximation, obtenu par filtrage du signal original. L'invariance à l'ajout d'une constante est une propriétés désirées en détection. Ainsi le filtre d'approximation $f_e(x)$ est :

1. passe-bande (composante continue nulle + localisation fréquentielle);
2. symétrique afin de ne pas modifier la phase du signal (donc pair).

L'équation (2.1) devient alors :

$$\begin{aligned} \hat{f}_A(x) &= f_e(x) * f(x) + i\mathcal{H}(f_e(x) * f(x)) \\ &= f_e(x) * f(x) + i\mathcal{H}(f_e(x)) * f(x) \\ &= (f_e(x) + if_o(x)) * f(x) , \end{aligned} \quad (2.4)$$

où $f_o(x)$ est la transformée de Hilbert de $f_e(x)$ et il est dit que les deux filtres sont en *quadrature*. On peut aisément montrer que $f_o(x)$ est un filtre passe-bande impair.

Ainsi, en pratique l'estimation est basée sur le calcul de la représentation analytique du signal d'approximation à une échelle donnée. Cette localisation en fréquence signifie que la phase comporte des informations sur les structures du signal à l'échelle d'analyse.

L'estimation utilise une paire de filtres en quadrature passe-bande (un filtre pair $f_e(x)$ et un filtre impair $f_o(x)$).

Le signal analytique et les filtres en quadrature correspondants fournissent un cadre très approprié pour l'extraction des caractéristiques locales des signaux (Knutsson et Andersson 2003b). Cet outil puissant est bien connu des praticiens en traitement du signal (Gabor 1946, Granlund et Knutsson 1995, Morrone et Owens 1987, Picinbono 1997).

2.1.2 Le cas multidimensionnel

Les notions de phase et d'amplitude locales sont intrinsèquement liées à la représentation analytique du signal. Cette dernière repose sur la transformée de Hilbert. L'application de ce signal pour le traitement d'images nécessite donc la généralisation de la transformée de Hilbert aux signaux multidimensionnels.

Une généralisation directe de celle-ci aux dimensions supérieures n'est pas évidente. Jusqu'à il y a une dizaine d'années, la théorie du traitement du signal multidimensionnel souffrait de l'absence d'une extension *isotrope* du signal analytique. La pratique était soit l'utilisation des filtres ré-orientables (Freeman et Adelson 1991, Simoncelli et Farid 1996) ou en effectuant l'analyse sur un ensemble d'orientations, puis en combinant simplement les résultats pour avoir une estimation pour une orientation donnée (Heitger 1995, Kovessi 1999, Morrone et al. 1995, Mulet-Parada et Noble 2000, Rejsfeld 1996, Robbins et Owens 1997, Venkatesh et Owens 1990).

Felsberg et Sommer (2001a) ont été les premiers à proposer un signal analytique isotrope 2D, *le signal monogénique*. L'idée de base est de concevoir un filtre isotrope impair vectoriel est non pas scalaire. Cette généralisation, qui en fait utilise la transformée de Riesz, à la place de la transformée de Hilbert, conserve les propriétés de base du signal analytique 1D qui décompose un signal en informations sur la structure (phase locale) et l'énergie (amplitude locale). De plus cette représentation nous permet l'estimation locale de l'orientation du signal. L'interprétation de la phase locale devient possible si la structure locale du signal est dite 'intrinsèquement 1D (voir eg. Wietzke et al. 2008).

L'isotropie de la construction a largement favorisé son utilisation pour résoudre divers problèmes de traitement d'image (Sommer et Zang 2007). Depuis des liens à la théorie des espaces échelles (Felsberg et al. 2005, Felsberg et Sommer 2004), et des extensions pour des images en couleur (Demarcq et al. 2011), aux ondelettes (Unser et al. 2009) et à des ordres supérieurs (Sedlazeck 2008, Unser et Van De Ville 2010, Wietzke et al. 2008, Wietzke et Sommer 2008) ont été proposées.

2.1.2.1 Détails pour le cas 2D

Dans le cas bidimensionnel, le signal monogénique d'un signal $f(x)$ est défini comme étant le triplet suivant :

$$f_M(x) = (f, h_x * f, h_y * f)(x), \quad (2.5)$$

obtenu par l'association du signal f et de sa transformée de Riesz :

$$f_R(x) = (h_x * f, h_y * f)(x) = h(x) * f(x) , \quad (2.6)$$

avec

$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} h_x \\ h_y \end{bmatrix} (x, y) = \frac{1}{2\pi(x^2 + y^2)^{\frac{3}{2}}} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (2.7)$$

ou d'une manière équivalente dans Fourier,

$$\mathbf{H}(\boldsymbol{\omega}) = \begin{bmatrix} H_x \\ H_y \end{bmatrix} (\omega_x, \omega_y) = \frac{-i}{\sqrt{\omega_x^2 + \omega_y^2}} \begin{bmatrix} \omega_x \\ \omega_y \end{bmatrix} = -i \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|}. \quad (2.8)$$

Il est simple de vérifier que cette représentation conserve les propriétés les plus intéressantes de la transformée de Hilbert 1D, notamment l'anti-symétrie et l'isotropie de sa fonction de transfert. Les notions d'amplitude, de phase et d'orientation locale sont définies d'une manière équivalente au cas 1D, en interprétant le signal monogénique dans un repère en coordonnées sphériques (Felsberg et Sommer 2000) :

$$\mathbf{f}_M(\mathbf{x}) = \begin{pmatrix} f(\mathbf{x}) \\ h_x * f(\mathbf{x}) \\ h_y * f(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} A_f(\mathbf{x}) \cos(\varphi(\mathbf{x})) \\ A_f(\mathbf{x}) \sin(\varphi(\mathbf{x})) \cos(\theta(\mathbf{x})) \\ A_f(\mathbf{x}) \sin(\varphi(\mathbf{x})) \sin(\theta(\mathbf{x})) \end{pmatrix}, \quad (2.9)$$

où $A_f(\mathbf{x}) = \sqrt{f(\mathbf{x})^2 + |f_R(\mathbf{x})|^2}$ représente l'amplitude, $\varphi(\mathbf{x}) \in [-\pi, \pi[$ est la phase et $\theta(\mathbf{x}) \in [0, \pi[$ est l'orientation dominante du signal au point \mathbf{x} . Notons que cette interprétation est justifiée plus amplement par la relation mise en évidence entre la transformée de Riesz avec la transformée de Radon et la transformée de Hilbert 1D (voir eg. Wietzke et Sommer 2008).

De la même manière que pour le 1D et pour les mêmes raisons, en pratique, c'est le signal monogénique d'une version filtrée du signal original qui est utile. Le filtre 2D d'approximation est l'extension isotropique du filtre pair passe-bande 1D $f_e(x)$. L'équation (2.4) devient alors en 2D :

$$\hat{\mathbf{f}}_M(\mathbf{x}) = (f_e * f, f_e * h_x * f, f_e * h_y * f)(\mathbf{x}). \quad (2.10)$$

Ainsi, le signal monogénique, à une échelle donnée, est obtenu par le filtrage du signal avec trois filtres passe-bande : Un premier filtre d'approximation, pair donc isotrope, et une paire de filtres impairs qui constitue dans leur ensemble le filtre impair isotrope. Un exemple des trois filtres dans le domaine de Fourier est illustré en figure 2.2.

2.2 Mesures de détection

Nous présenterons ci-dessous les deux mesures les plus populaires utilisées pour la détection de contours en utilisant l'information de phase locale.

2.2.1 La congruence de phase

La congruence de phase mesure le degré d'alignement local entre les composantes de Fourier. Elle peut être estimée en mesurant la déviation moyenne de la phase des composantes de Fourier. Étant donnée le développement en série de Fourier d'un signal

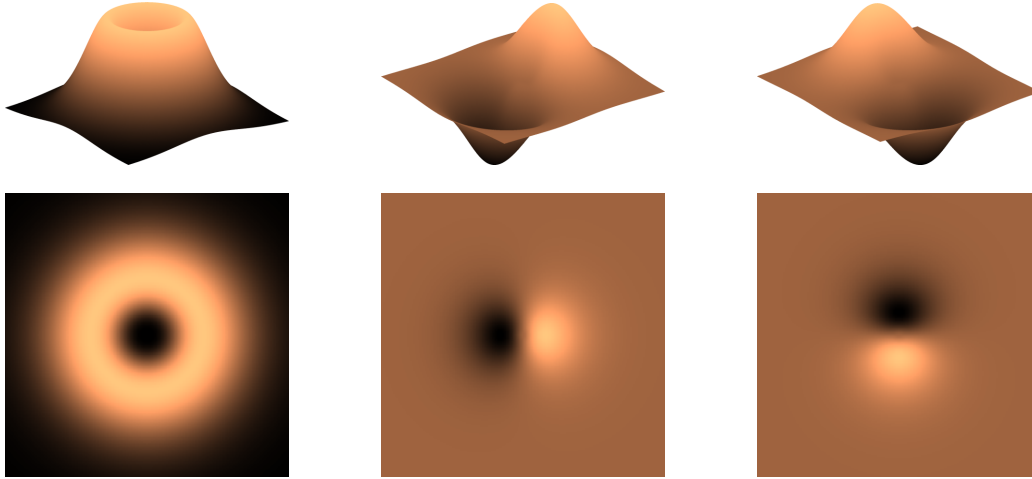


Figure 2.2 – Exemple de filtres passe-bande en quadrature 2D dans le domaine de Fourier. De gauche à droite : filtre pair isotrope $H_e(\omega)$, la paire de filtres impairs $(H_x(\omega), H_y(\omega))$. $H_e(\omega)$ est réel pur ; les deux autres sont imaginaires purs.

unidimensionnel, $f(x) = \sum_n A_n \cos(\varphi_n(x))$, où $A_n(x)$ et $\varphi_n(x)$ représentent respectivement l’amplitude et la phase de la $n^{ième}$ composante de Fourier en un point x . Alors la congruence de phase du signal $f(x)$ est définie par [Morrone et Owens \(1987\)](#) :

$$PC(x) = \max_{\bar{\varphi} \in [0, 2\pi[} \frac{\sum_n A_n(x) \cos(\varphi_n(x) - \bar{\varphi})}{\sum_n A_n(x)} . \quad (2.11)$$

[Venkatesh et Owens \(1990\)](#) ont montré que l’énergie locale d’un signal est égale à la congruence de phase pondérée par la somme des amplitudes de Fourier.

$$A(x) = PC(x) \sum_n A_n(x) . \quad (2.12)$$

Par conséquent, la recherche des maxima de l’énergie locale est équivalent à la recherche des points là où la congruence de phase est maximale. Cette équation possède une interprétation géométrique très simple. En effet, si l’on considère les composantes de Fourier comme des vecteurs de longueur $A_n(x)$ et d’angle $\varphi_n(x)$ alors l’énergie locale est à son maximum lorsque toutes les composantes de Fourier sont localement en phase. Cette dernière équation a servi de point de départ pour [Kovesi \(1999\)](#) pour définir plusieurs estimateurs de PC en utilisant un banc de n filtres passe-bande en quadrature.

2.2.2 Feature Asymmetry

[Kovesi \(1996\)](#) a proposé un nouvel outil de détection de contours de type bord (ou ligne). Le principe repose sur une mesure d’asymétrie (symétrie) basée sur l’information de phase locale. En effet, pour détecter des structures de type bord par exemple, cela revient à trouver les points où la phase locale $\varphi = \pm \frac{\pi}{2}$. En d’autres termes, on recherche les points où la réponse du filtre impair en valeur absolue est maximale et celle du filtre pair est minimale (i.e. en terme d’énergie). Kovesi a proposé le calcul d’une mesure, appelée “Feature Asymmetry” définie en utilisant la différence entre les valeurs absolues

des réponses de N filtres impairs et pairs (Kovesi 1997) :

$$FA(x) = \frac{\sum_n \lfloor A_n [|\sin(\varphi_n(x))| - |\cos(\varphi_n(x))|] - t_n \rfloor}{\sum_n A_n + \varepsilon}, \quad (2.13)$$

$$= \frac{\sum_n \lfloor |o_n(x)| - |e_n(x)| - t_n \rfloor}{\sum_n A_n(x) + \varepsilon}, \quad (2.14)$$

où pour une échelle n donnée, o_n et e_n représentent respectivement les réponses des filtres impair et pair, $A_n = \sqrt{o_n^2 + e_n^2}$ représente l'amplitude, $\lfloor \cdot \rfloor$ représente l'opérateur de troncature des valeurs négatives à zéro et t_n est un seuil de bruit (pour l'estimation de t_n , voir eg. Kovesi 1999, Mulet-Parada et Noble 2000). Cette mesure prend ses valeurs dans l'intervalle $[0, 1]$; elle est très proche de 1 au voisinage des contours.

Notons que certains auteurs ont utilisé cette mesure avec une seule échelle (Boukerroui et al. 2001, Mulet-Parada et Noble 2000), et d'autre l'ont définie comme étant la somme des rapports et non pas le rapport des deux sommes (Belaid et al. 2011, Rajpoot et al. 2009) :

$$FA = \frac{1}{N} \sum_n \frac{\lfloor |o_n| - |e_n| - t_n \rfloor}{A_n + \varepsilon}. \quad (2.15)$$

La mesure de symétrie se déduit facilement en faisant un raisonnement similaire en s'intéressant, cette fois-ci, aux points où la phase locale $\varphi \in \{0, \pi\}$.

2.3 Contributions

Dans le contexte d'extractions de caractéristiques, bien que l'importance des filtres en quadrature a été mise en évidence (Knutsson et Andersson 2003b), malheureusement, peu de travail a été fait sur la sélection de la partie radiale des filtres en quadrature. La plupart des publications utilisaient la famille des log-Gabor, principalement pour avoir une composante continue nulle. Plusieurs familles de paires de filtres en quadrature ont été proposées et appliquées dans la littérature : Gabor, log-Gabor, dérivées de gaussiennes, différence de gaussiennes (DoG), Cauchy, etc. Nous présentons ci-dessous nos contributions.

2.3.1 Synthèse et analyse de l'existant

J'ai initié ce travail à la fin de mon post-doc à Oxford. À l'époque, le papier de Felsberg et Sommer (2001a) venait juste de paraître.

Nous avons fait un travail de synthèse en recensant et étudiant la majorité des familles de filtres en quadrature proposées dans la littérature. Nous avons aussi proposé la construction de nouvelles paires de filtres en quadrature issues des opérateurs de détection optimaux au sens des critères de Canny (1986) :

1. Détecteur de contours de Deriche (1987)
2. Détecteur de contours de Sarkar et Boyer (1991)
3. Détecteur de lignes de Ziou (1991),

Dans notre étude nous avons choisi de travailler sur la représentation fréquentielle des filtres car certaines familles ne possèdent pas une expression analytique des noyaux

de convolution. Cela nous a permis de généraliser certaines familles en nous ne limitant pas à des ordres de dérivée entiers.

L'étude portait premièrement sur la paramétrisation des différentes familles en terme de fréquence centrale (échelle d'analyse) et de bande passante. Deux caractéristiques très utiles pour la mise en œuvre pratique, mais aussi nécessaires pour une comparaison objective. Dans la mesure du possible, une comparaison au sens des critères de Canny est fournie. Nous avons également étudié la stabilité en phase des différentes familles, comme suggéré par Fleet et Jepson (1993). Le comportement des filtres en terme de localisation spatiale et fréquentielle est également analysé via le principe d'incertitude. Nous nous sommes aussi intéressés aux corrections de la composante continue des filtres de Gabor. Deux types de correction existent dans la littérature (Boukerroui et al. 2006). Une proposée par Ronse (1993) et une seconde avec deux variantes, connues sous le nom de la famille des S-Gabor (Stretched-Gabor) (Heitger 1995, Heitger et al. 1992). L'introduction de la correction peut dévier considérablement les propriétés du filtre. En effet, cela s'accompagne par un décalage de la fréquence centrale, une perturbation de la bande passante et une énergie non nulle dans la partie des fréquences négatives, indiquant que la paire de filtres (paire et impaire) n'est plus en quadrature. Les deux types de corrections sont analysés et comparés. De plus une variante de la correction de Ronse est proposée, limitant considérablement ses distorsions.

Le bilan de ce travail est une mise à plat, pour la première fois dans littérature à notre connaissance¹, de toutes les familles de filtres en quadrature, utilisées dans des problématiques de détection ou d'estimation de mouvement. Nous en avons également proposé de nouvelles et généralisé d'autres. Les propriétés principales des familles étudiées sont résumées dans un tableau dans [5], disponible en page (119). Dans le contexte de la détection, il ressort de notre analyse que les filtres de Gabor sont un mauvais choix. Et c'est probablement le cas aussi pour la famille des log-Gabor, largement utilisée dans la littérature. Les familles des DoGs et Cauchy ont de meilleures propriétés. Aussi, il faut généralement choisir des filtres à large bande (>2 octaves) pour une meilleure détection. Nous préconisons également l'utilisation des dérivées d'une gaussienne ou les filtres de Cauchy pour la problématique d'estimation de mouvement avec un algorithme de phase locale.

Enfin, précisons que cette étude n'inclut pas certaines familles de filtres, notamment certaines familles apparues récemment (Felsberg et Sommer 2004, Knutsson et Andersson 2003a, Mellor et al. 2008) et les filtres de Simoncelli et al. (1992) car l'avantage principal, la propriété de "shiftabilité" de la famille, impose leur utilisation dans une approche multi-échelles.

2.3.2 Nouvelles familles de filtres

C'est dans le double contexte de la théorie des espaces d'échelles linéaires et des filtres en quadrature que Felsberg et Sommer (2004) ont introduit l'espace d'échelles de Poisson monogénique. Cette théorie a ouvert de nouvelles possibilités de traitement d'images en utilisant l'information de la phase locale. Initialement, Felsberg et Duits ont travaillé indépendamment sur l'espace des échelles de Poisson (Duits et al. 2004, Felsberg et Sommer 2001b), puis, il ont travaillé ensemble (Duits et al. 2003, Felsberg et al.

1. En fait, Perona et Malik (1990a) ont travaillé sur la même problématique des choix des filtres en développant une approche similaire à celle de Canny (1986). Puis un travail est apparu sur les espaces des échelles des filtres en quadrature dans (Kube et Perona 1996). Nous n'avons pas connaissance de ces travaux avant 2008.

2005). Duits et al. (2004) ont travaillé sur les axiomes que vérifie une forme généralisée des espaces des échelles linéaires, définie pour la première fois dans (Pauwels et al. 1995). Il s'agit d'une classe paramétrée par $\alpha \in]0, 1]$ qui inclut l'espace des échelles de Poisson ($\alpha = 1/2$) ainsi que le populaire gaussien ($\alpha = 1$) (Witkin 1983). Une récente extension au cas non-linéaire est donnée dans (Feigin et al. 2010). Nous recommandons au lecteur de se référer à (Lindeberg 2011) pour un article récent sur l'espace des échelles gaussien et à (Sommer et Zang 2007) pour une vue d'ensemble sur l'espace des échelles monogénique des signaux multidimensionnels. Enfin notons que la famille des filtres en quadrature, connue auparavant par le nom de Cauchy n'est finalement que celle des dérivées du noyau de Poisson.

Récemment, nous avons fait usage de cette nouvelle représentation unifiée pour dériver deux nouvelles familles de filtres en quadrature passe-bande, construites à partir de dérivées ou de la différence de deux noyaux générateurs. Ces nouveaux filtres incluent les familles de filtres communément issues des noyaux gaussien et de celui de Poisson.

Les propriétés de chaque famille sont d'abord étudiées en suivant une démarche similaire à notre étude précédente. Dans un second temps, nous avons examiné expérimentalement l'influence du paramètre α dans le contexte de détection de points de contour en utilisant un critère de détection basé sur la phase locale, sur des simulations d'images ultrasonores. Ces résultats préliminaires suggèrent que le paramètre optimal α dépend de la famille de filtres (différence ou dérivée) et peut être ajusté expérimentalement. Ce travail va bientôt paraître dans la conférence IEEE ISBI [12], et dans un article du journal Pattern Recognition [1].

Publications : [5] (cf. page 103), [12] (cf. page 135), [11], [1].

REHAUSSEMENT DE CONTRASTE EN ÉCHOCARDIOGRAPHIE

3.1 Introduction

Une manière classique d'améliorer les performances des détecteurs de points de contours, est de les faire précéder par une étape de débruitage. Ce dernier est un traitement important, largement étudié et appliqué (pas seulement dans le contexte de segmentation d'images). Cependant, dans de nombreuses applications, la réduction du bruit n'est qu'une solution partielle car elle ne peut pas corriger l'image des artefacts introduits par le système d'imagerie. Plus particulièrement, en imagerie médicale, les artefacts dépendent souvent de l'objet imagé, ce qui conduit finalement à des régions non homogènes dans l'image. Ce problème est bien connu par les traiteurs d'images médicales. L'inhomogénéité d'intensité touche particulièrement deux modalités d'image : l'IRM et l'échographie B-scans. Il est particulièrement bien étudié en IRM (voir eg. [Vovk et al. 2007](#)) et très utilisé comme prétraitement en segmentation ([Gordillo et al. 2013](#)). En revanche, cette problématique a reçu relativement peu d'attention en imagerie ultrasonore, en raison du très faible rapport signal à bruit qui caractérise cette modalité. L'inhomogénéité est essentiellement due au phénomène de l'atténuation des ondes acoustiques en déplacement dans le tissu imagé et à des artefacts tels que le renforcement postérieur, l'ombre acoustique ou la réflexion oblique.

Durant mon séjour post-doctoral, mes co-auteurs et moi avons proposé une méthode originale pour traiter ce problème sur des séquences d'images échocardiographiques [26, 8]. Notre approche s'appuie sur un algorithme de détection de contours basé sur la phase locale, suivi par une méthode d'interpolation de surface à partir d'un nuage de points et un post-traitement non-linéaire en fin de chaîne de traitement. Nous exploitons d'abord la propriété d'invariance au contraste de l'image de notre détecteur afin de sélectionner un ensemble de caractéristiques pertinentes de l'image. Ensuite, une approximation basse fréquence des caractéristiques sélectionnées est obtenue en utilisant un algorithme d'interpolation de surface rapide. Enfin, un post-traitement non-linéaire est appliqué pour obtenir l'image corrigée. Notre approche repose sur *l'a priori* que les caractéristiques sélectionnées devraient avoir des niveaux de gris comparables sur l'image après correction.

Nous allons résumer ci-dessous notre travail en détaillant principalement la méthode d'interpolation de surface.

3.2 Interpolation de surface

L'interpolation de surface à partir d'un ensemble de points de mesures bruitées est un problème mal posé : un ensemble infini de surfaces peut satisfaire un ensemble donné de contraintes. Par conséquent, une procédure de régularisation, en tenant compte de la pertinence visuelle de la solution et de l'efficacité de calcul est généralement appliquée, de sorte que le problème d'interpolation devient celui de la minimisation d'une fonctionnelle d'énergie de la forme :

$$U(f) = U_d(f, d) + \lambda U_r(f) . \quad (3.1)$$

Le premier terme mesure la fidélité de la surface d'approximation f aux données mesurées d . Le second, est le terme de régularisation modélisant notre *a priori* sur la solution recherchée avec λ un paramètre réel positif de contrôle. Une attache aux données couramment utilisée est la somme des erreurs quadratiques pondérées entre le champs d'observations $d = \{d_i, i = 1 \dots n\}$ aux points x_i et la fonction d'approximation $f(x_i)$,

$$U_d(f, d) = \sum_{i=1}^n w_i (f(x_i) - d_i)^2 , \quad (3.2)$$

où $0 \leq w_i \leq 1$ sont des pondérations qui reflètent un degré de confiance (mesuré ou *a priori*) sur les observations d_i . Parmi les différents termes de régularisation possibles, nous avons considéré l'énergie d'une plaque mince (Thin plate spline), puisqu'*a priori* notre surface ne devrait pas avoir de discontinuités. En deux dimensions, cela donne :

$$U_r(f) = \int \int \left[\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right] dx dy . \quad (3.3)$$

L'approximation du problème continu en utilisant des opérateurs discrets permet une résolution numérique. En supposant que les données sont définies sur une grille régulière $S = \{x_i, y_i, 1 \leq i, j \leq N\}$, la discrétisation de la surface sur la grille S nous conduit à l'ensemble de variables nodales $v = \{f(x_i, y_i), 1 \leq i, j \leq N\}$ alors l'utilisation des différences finies pour l'approximation des dérivées secondes dans (3.3) nous permet d'écrire l'énergie (3.1) sous la forme matricielle suivante :

$$U(v) = v^T A v - 2v^T b + c , \quad (3.4)$$

où la matrice $A = A_w + \lambda A_r$ et le vecteur $b = A_w d$ et c est une constante. La matrice A_r est issue de l'écriture matricielle de la discrétisation de (3.3) et $A_w = \text{diag}\{w_{ij}\}$. Ainsi la solution est donnée par la résolution du système linéaire $A v = b$ avec une très grande matrice creuse de taille $N^2 \times N^2$.

3.2.1 Résolution dans l'espace ondelettes

Le problème d'interpolation de surface décrit dans la section précédente conduit à la résolution d'un très grand système linéaire avec une matrice très creuse. Les caractéristiques de la matrice nous permettent de croire qu'elle a un mauvais conditionnement. Par conséquent, l'utilisation des méthodes itératives simples conduisent à une convergence très faible. Classiquement, les méthodes multi-grilles sont utilisées pour accélérer la convergence. Pour notre part, nous avons utilisée l'approche de [Yaou et Chang \(1994\)](#).

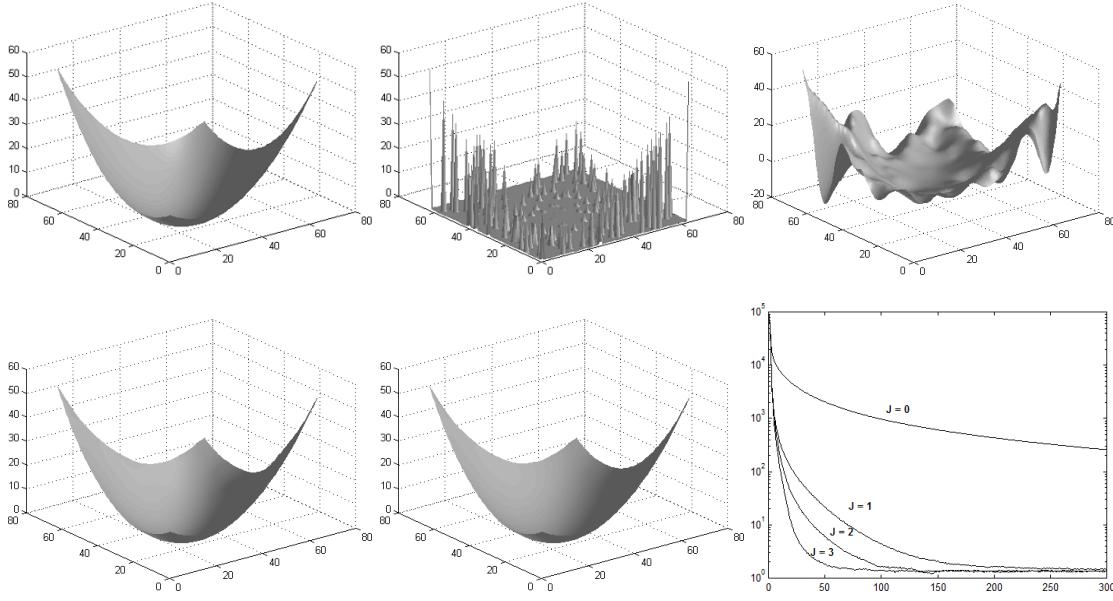


Figure 3.1 – Exemples de résultats de l’interpolation de surface par la TOD. De gauche à droite ; 1^{ère} ligne : surface de référence ; données observées (5% de la surface de référence) ; résultat sans preconditionnement ($J = 0$). 2^{ème} ligne : résultats avec preconditionnement pour $J = 2$ puis $J = 3$ niveaux de décompositions de la TOD. Évolutions de l’énergie (3.4) en fonction des itération. Les résultats sont montrés après 300 itérations de la méthode de Gauss-Seidel. $\lambda = 0.1$ et $v_0 = d$.

Cette approche utilise le concept de preconditionnement via la transformée en ondelettes discrète (TOD).

Les auteurs exploitent le fait que généralement les composantes hautes fréquences du problème d’interpolation convergent plus rapidement que les basses fréquences. Il est intéressant de noter que c’est le même argument qui a inspiré l’utilisation des techniques multi-grilles. L’utilisation de la TOD permet une résolution séparée des basses fréquences des hautes fréquences. Pratiquement, au lieu de rechercher directement v on recherche sa décomposition sur une bases d’ondelettes (B-spline) $\tilde{v} = Dv$ avec D la matrice de décomposition de la TOD et $v = R\tilde{v}$ où R est la matrice de reconstruction. Ainsi, on peut montrer que la minimisation de l’énergie (3.4) exprimée avec \tilde{v} conduit au système linéaire suivant :

$$\tilde{A}\tilde{v} = \tilde{b} \quad \text{avec} \quad \tilde{A} = R^T A R \quad \text{et} \quad \tilde{b} = R^T b \quad (3.5)$$

Un exemple de résultat sur une surface de synthèse, avec seulement 5% de données observées, montre l’efficacité de l’approche. Notons cependant que le calcul de la matrice \tilde{A} nécessite beaucoup d’effort tant en mémoire qu’en temps de calcul.

3.3 Détection de contours

Dans ce travail, nous avons opté pour la mesure de *Feature Asymmetry* utilisée par Mulet-Parada et Noble (2000). Le principe de la mesure est détaillée précédemment (§ 2.2.2 page 46). Cette mesure offre une bonne détection des discontinuités asymétriques. De plus, elle a l’avantage d’être théoriquement invariante au contraste de l’image. Aussi la mesure est normalisée entre 0 et 1 et sa valeur peut être interprétée

comme une mesure de confiance sur la détection. Cela nous a permis de définir l'attachement aux données nécessaire pour l'étape d'interpolation de surface comme suit

$$d = \{d_{ij} = \max_{B_{ij}} I(x, y) \quad \text{si } FA(x_i, y_j) > 0; \quad 1 \leq i, j \leq N\} \quad (3.6)$$

$$w = \{w_{ij} = FA(x_i, y_j); \quad 1 \leq i, j \leq N\} , \quad (3.7)$$

où I représente l'image originale et B_{ij} est une petite fenêtre rectangulaire centrée sur le pixel (x_i, y_j) . Prendre la valeur maximale de l'intensité de l'image dans une fenêtre centrée sur la position du point de contour détecté garantit que l'on prend toujours la valeur la plus élevée de l'arête. Par conséquent, l'approche est plus robuste aux erreurs de localisation. Notons que l'orientation du contour peut éventuellement être prise en compte lors de la définition du voisinage.

3.4 La méthode de correction

Le modèle physique de la formation des images ultrasonore suggère un modèle multiplicatif de la dégradation basse fréquence. Motivé par cela, nous avons proposé l'équation de correction suivante

$$I_c(x, y) = \frac{I(x, y) / \max(I(x, y))}{v^*(x, y) / \max(v^*(x, y)) + \gamma} , \quad (3.8)$$

où $v^*(x, y)$ est la surface interpolée et γ est un paramètre positif de contrôle qui assure que $I_c(x, y) \propto I(x, y)$ si $\gamma \gg 1$. La correction maximale est obtenue pour $\gamma \ll 1$.

3.5 Résultats et bilan

Afin de tester et d'évaluer quantitativement les performances de notre méthode de rehaussement sur des séquences d'images échocardiographiques et d'étudier l'influence du paramètre de contrôle γ nous avons proposé deux mesures :

1. Une mesure de la hauteur des arrêtes de l'endocarde et de l'épicarde.
2. Deux mesures de la consistance de la correction dans le temps.

La méthode est testée sur plusieurs séquences d'images réelles et évaluée quantitativement sur trois séquences. Nous l'avons aussi testée sur une séquence d'images de synthèse afin de vérifier que la méthode n'introduit pas d'artefact sur les images corrigées. Le paramètre de contrôle γ est facilement réglable. Pour notre application, une valeur autour de 0.2 offre un bon compromis entre le rehaussement des régions sombres et très claires de l'image. Quelques images avant et après correction sont montrées en figures 3.2 et 3.3. On peut apprécier l'amélioration de la qualité des images et de la détection des contours obtenue sur les images après correction. Notons enfin que le but de notre approche n'est pas de filtrer le bruit du speckle mais seulement de réduire l'inhomogénéité de l'image. Cela permet de préserver l'information utile contenue dans la texture de ce bruit. Ainsi, cela facilite une interprétation par un expert médical de l'image et permet aussi l'utilisation en aval d'une modélisation statistique du bruit dans un traitement numérique de l'image corrigée.

Publications : [8] (cf. page 141), [26].

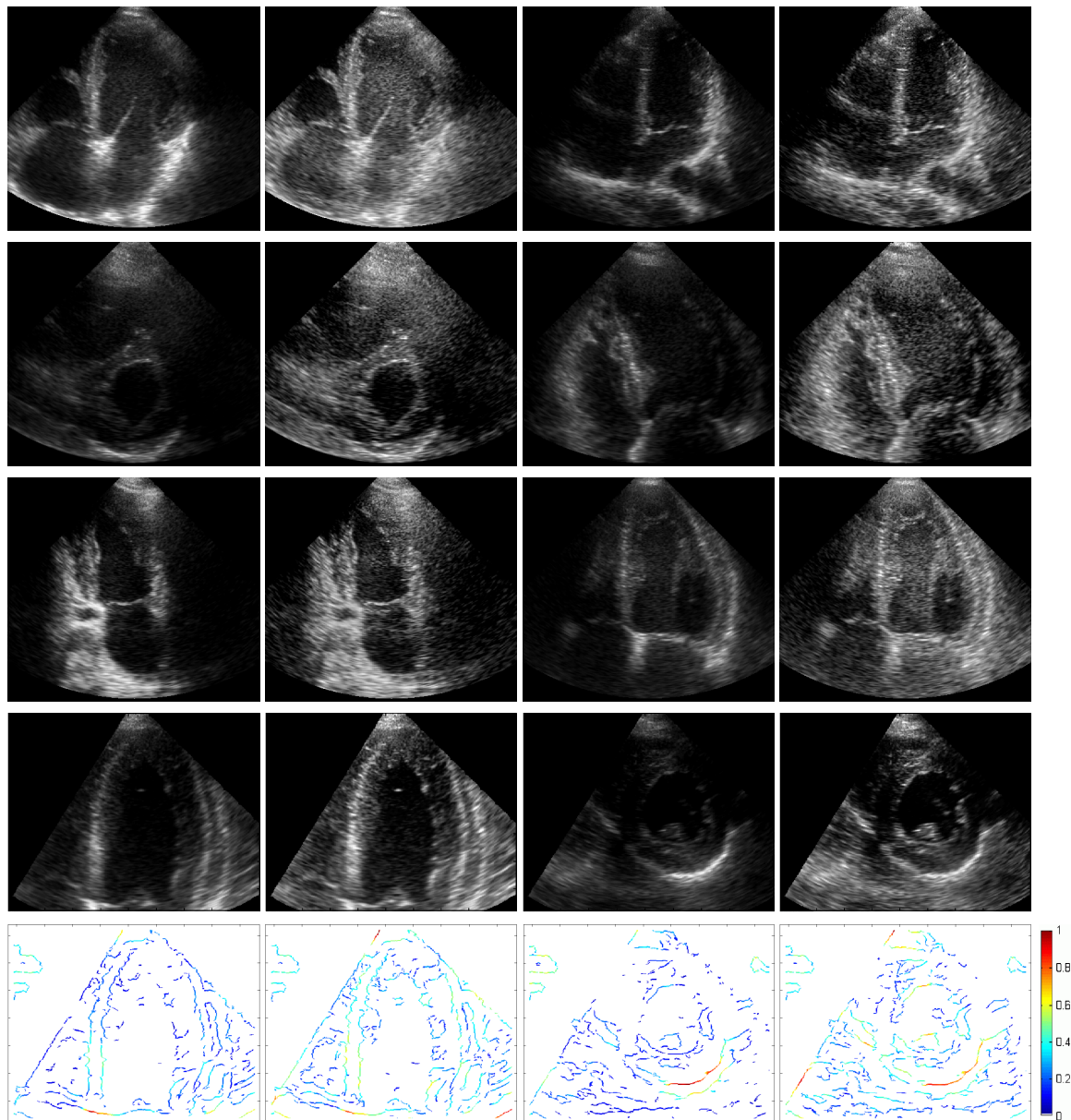


Figure 3.2 – Exemples de résultat de correction obtenues pour $\gamma = 0.2$. De gauche à droite : image originale ; image corrigée ; image originale ; image corrigée. La dernière ligne montre les résultats de la mesure de détection *FA* sur les images de la ligne précédente.

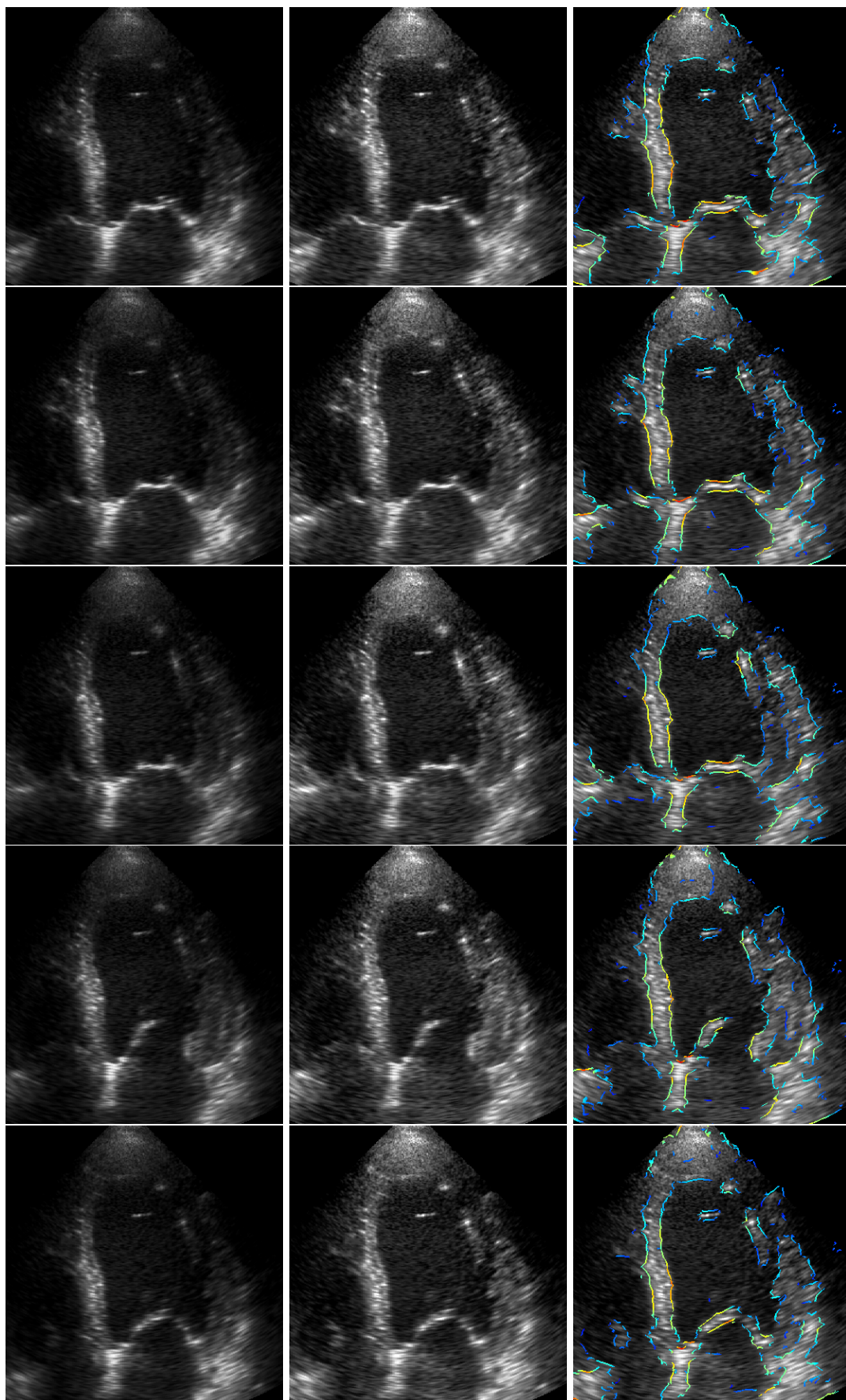


Figure 3.3 – Images n° 1, 4, 8, 11, et 14 d’une même séquence. De gauche à droite : image originale ; corrigée et corrigée avec les contours détectés.

ESTIMATION DE MOUVEMENT EN ÉCHOGRAPHIE

L'estimation du flot optique ou du mouvement est un problème fondamental en vision par ordinateur. La littérature sur le sujet est très abondante et le problème reste toujours d'actualité (Baker et al. 2011, Kondermann et al. 2012, Sun et al. 2013). La précision des algorithmes d'estimation progresse régulièrement comme en témoignent les résultats sur le *benchmark* de Middlebury (Baker et al. 2007)¹. Cette problématique occupe également une place importante en imagerie médicale et plus particulièrement en échographie (Coatrieux 2005, 2006). En effet, cette modalité d'image étant de type temps réel, elle est particulièrement adaptée pour l'étude des tissus ou des fluides en mouvement ou en déformation, comme c'est le cas en élastographie. Notant que l'IRM progresse à très grand pas dans ce champ d'application et plus récemment encore la tomodographie à rayon X multi-coupes (Multi-Slice Computed Tomography). De nos jours, ces modalités d'images sont très utilisées dans le diagnostic et l'étude de la fonction cardiaque (Tobon-Gomez et al. 2013, Wang et Amini 2012). Néanmoins, l'échocardiographie est de loin la plus utilisée en clinique.

Au delà du formalisme utilisé pour résoudre le problème d'estimation du mouvement, les techniques appliquées en imagerie médicale peuvent être classées comme appartenant à l'un des trois principaux groupes suivants : les techniques différentielles ou basées sur les gradients spatiaux et temporels s'appuient sur l'hypothèse que la luminosité d'un motif est invariante dans le temps. Dans la deuxième catégorie on trouve les méthodes formalisées dans le domaine de Fourier. Elles s'appuient sur l'estimation de la phase locale par un banc de filtres linéaires en quadrature. La phase étant théoriquement indépendante du contraste de l'image, cela suggère leur adéquation à la problématique d'estimation de mouvement sur des images échocardiographiques². La troisième classe regroupe des méthodes du type mise en correspondance de blocs (block-matching). Le mouvement étant calculé comme les déplacements de ces blocs et estimé en maximisant/minimisant une mesure de similarité/dissimilarité. Cette catégorie s'apprête donc mieux à la prise en compte de la nature statistique du bruit dans le processus de formation de l'image. Les trois catégories supposent, d'une manière implicite ou explicite, que le mouvement est similaire pour tous les pixels dans un bloc compact au voisinage d'un pixel donné.

1. <http://vision.middlebury.edu/>.

2. Voir les travaux récents de Alessandrini et al. (2013a,b)

De ce qui précède, il est clair que la validité des hypothèses émises dans l'estimation de mouvement peut être au moins remise en question. L'hypothèse de régularité du mouvement n'est souvent pas vérifiée et l'hypothèse de conservation de l'intensité le long des trajectoires est en général violée. Un éclairage sur les raisons de non respect de celles-ci dans le contexte spécifique de l'imagerie médicale et de leur incidence sur la qualité et la robustesse de l'estimation est donné par [Coatrieux \(2006\)](#). Toutefois, il apparaît que les performances de l'ensemble de ces méthodes ont été grandement améliorées récemment par couplage entre les méthodes de flot optique et les méthodes de correspondance de blocs. Notons que ces dernières sont plus particulièrement adaptées lorsque l'estimation du mouvement n'est pas requise en tout point de l'image. C'est dans ce contexte précis que s'inscrit mon travail effectué en majeure partie durant mon stage postdoctoral à Oxford.

Motivés par les résultats obtenus dans ([Mikić et al. 1998](#)) et par une mesure de similarité originale – et récente à l'époque – dérivée en fonction d'un modèle de la formation des images ultrasonores simplifié ([Cohen et Dinstein 2000, 2002](#)), nous avons développé une nouvelle méthode du type block-matching pour le suivi d'un contour fermé (mais pas nécessairement) sur une séquence d'images ultrasonores. La méthode d'estimation proposée s'appuie sur l'algorithme de [Singh et Allen \(1992\)](#). Notons que notre objectif principal est l'application de l'algorithme pour le suivi de l'endocarde sur des séquences d'images échocardiographiques.

Nous allons résumer ci-dessous notre travail.

4.1 Mesure de similarité

Un des modèles les plus simples du speckle issu de la modélisation de l'interaction physique de l'onde ultrasonore avec le tissu imagé est celui du bruit multiplicatif avec une densité de probabilité de Rayleigh. Le speckle est dit alors *pleinement développé* car le modèle suppose un nombre de diffuseurs dans le milieu important. Rappelons que ce bruit caractérise l'enveloppe du signal radio-fréquence rétro-diffusé avant tout autre post-traitement (interpolation, compression logarithmique...). [Cohen et Dinstein \(2002\)](#) ont utilisé ce modèle pour dériver une mesure de similarité entre deux blocs de pixels x et y . Si l'on suppose que les deux blocs sont deux observations indépendantes d'un même tissu s , alors

$$x_j = \eta_j^1 s_j, \quad y_j = \eta_j^2 s_j, \quad (4.1)$$

avec j un indice parcourant les pixels dans le bloc. η_j^1 et η_j^2 sont les réalisations, supposés iid (indépendantes et identiquement distribuées), de deux bruits de type Rayleigh de paramètre α^2 et β^2 respectivement. De l'équation précédente nous pouvons écrire :

$$x_j = \eta_j y_j, \quad (4.2)$$

avec $\eta_j = \eta_j^1 / \eta_j^2$. L'application de la fonction logarithme au modèle multiplicatif précédent nous conduit au modèle de bruit additif suivant

$$\tilde{x}_j = \tilde{y}_j + \tilde{\eta}_j \quad \text{avec} \quad p(\tilde{\eta}) = 2 \left(\frac{\alpha}{\beta} \right)^2 \frac{\exp(2\tilde{\eta})}{\left(\exp(2\tilde{\eta}) + \left(\frac{\alpha}{\beta} \right)^2 \right)^2}, \quad (4.3)$$

où la notation \tilde{x} désigne les données transformées par la fonction log. L'estimation du maximum de vraisemblance du mouvement, dérivée à partir du modèle de bruit ci-

dessus et supposant que le bruit observé dans les deux blocs est iid (c-à-d $\alpha = \beta$), nous conduit à la mesure de similarité définie dans (Cohen et Dinstein 2002) :

$$E_{CD_2} = \sum_{j=1}^{|\tilde{x}|} \{ \tilde{x}_j - \tilde{y}_j - \ln(\exp(2(\tilde{x}_j - \tilde{y}_j)) + 1) \} . \quad (4.4)$$

4.2 L'approche de Singh

L'approche de Singh de mise en correspondance de blocs s'appuie essentiellement sur la définition d'une fonction de densité de probabilité (fdp) du mouvement en chaque pixel à partir de la mesure de similarité/dissimilarité entre le bloc centré autour du pixel considéré et tous les blocs dans la fenêtre de recherche. Une première estimation du vecteur de déplacement, $v_c = (u_c, v_c)^T$, en chaque point de l'image, est alors donnée par la moyenne de cette densité. Une seconde estimation est obtenue en supposant qu'*a priori* des pixels voisins doivent avoir des vecteurs déplacements similaires. Formellement, cet *a priori* est modélisé par une probabilité gaussienne autour du voisinage de chaque pixel. Ainsi, une seconde estimation, $v_n = (u_n, v_n)^T$, peut-être définie comme étant la moyenne pondérée des déplacements des pixels voisins. Les deux estimations étant issues d'un modèle probabiliste, les erreurs d'estimation peuvent être quantifiées par leurs matrices de covariances respectives S_c et S_n . Singh et Allen (1992) ont défini comme estimation optimale du vecteur de mouvement en chaque pixel, celui qui minimise l'erreur totale. Cette dernière est définie comme égale à la somme de l'erreur du terme d'attache aux données avec celle du terme de régularisation :

$$\epsilon^2(v) = (v - v_c)^T S_c^{-1} (v - v_c) + (v - v_n)^T S_n^{-1} (v - v_n) . \quad (4.5)$$

Étant donné que le terme de régularisation (voisinage) est défini en fonction de l'estimation courante, en pratique le problème de minimisation est résolu de manière itérative :

$$\begin{cases} \hat{v}^0 &= v_c , \\ \hat{v}^{m+1} &= \left[S_c^{-1} + (S_n^m)^{-1} \right]^{-1} \left[S_c^{-1} v_c + (S_n^m)^{-1} \hat{v}_n^m \right] . \end{cases} \quad (4.6)$$

4.3 Contributions

La mesure de similarité CD2, donnée en (4.4), est certes mieux adaptée pour les images ultrasonores car elle prend en compte la nature multiplicative du bruit. De plus, elle intègre la compression logarithmique de l'image enveloppe et la caractérisation statistique du bruit. Cependant, le modèle suppose que la distribution du bruit dans chacun des deux blocs est la même. Il y a de multiples raisons de croire à la non-validité de l'hypothèse sur des données acquises en clinique. Pratiquement, nous avons normalisé l'intensité des deux blocs pour avoir une moyenne nulle et une variance unité avant le calcul de la mesure de similarité. Notons que cette démarche n'a aucune justification théorique. Cependant, nos expérimentations ont montré qu'elle améliore grandement les résultats par rapport à une estimation des paramètres α et β .

Notre contribution majeure dans ce travail réside dans la paramétrisation de la fdp du terme d'attache aux données. En effet, il apparaît clairement que le compromis entre les deux termes qui constituent l'équation (4.5) dépend des matrices de covariances.

Singh et Allen (1992) n'ont pas normalisé la mesure de similarité utilisée (SSD), ce qui a pour conséquence d'engendrer des difficultés de paramétrisation de la fdp car cette dernière dépend de la taille du bloc utilisé. Pour notre part, nous l'avons définie comme suit :

$$R_c(u, v) = \frac{1}{Z} \exp \left\{ k \frac{E_c(u, v) - m}{(2n + 1)^2} \right\} \quad - N \leq u, v \leq N, \quad (4.7)$$

où E_c la mesure de similarité, m est son maximum dans la fenêtre de recherche W_s de taille $(2n + 1)^2$, et $(2n + 1)^2$ est la taille du bloc. Une façon d'obtenir des estimations plus robustes de la vitesse est de biaiser l'estimation vers le mode dominant de la densité. Une façon de faire est de prendre comme estimation la moyenne d'une version seuillée de la fdp

$$R_c^h(u, v) = \begin{cases} R_c(u, v) & \text{si } R_c(u, v) \geq \alpha(h), \\ 0 & \text{sinon ;} \end{cases} \quad (4.8)$$

où le seuil $\alpha(h)$ est défini comme $\alpha(h) = \widehat{m} - h(\widehat{m} - \widetilde{m})$ avec $h \in [0, 1]$. \widehat{m} et \widetilde{m} étant le maximum et le minimum de la fdp R_c dans la fenêtre W_s . Il est important de noter que la matrice de covariance S_c est calculée en utilisant la fdp non seuillée. Par conséquent, l'erreur du terme d'attache aux données prend en compte la présence d'un second mode. Le couple de paramètres de contrôle (k, h) peut être optimisé en minimisant une erreur de recalage sur une séquence d'images.

Enfin, pour notre application de suivi de contours, le terme de régularisation est estimé sur un voisinage le long du contour.

4.4 Résultats et bilan

La paramétrisation s'est avérée très utile en pratique. En effet, pour $h = 0$ l'estimation correspondrait à prendre le déplacement qui maximise la mesure de similarité, ou le mode de la fdp. Nos résultats montrent que c'est le pire des scénarios. Notre approche est équivalente à celle de Singh pour $h = 1$. Dans ce cas, l'optimisation sur le paramètre k améliore les résultats d'environ 15%. Une amélioration de 35% est observée par l'optimisation du couple de paramètres (voir figure page 164). Un avantage non négligeable, d'un point de vue applicatif, est la rapidité de l'approche.

Ce travail est très cité dans la littérature. Comme nous l'avons mentionné dans l'introduction, le couplage entre les méthodes de flot optique et les méthodes de correspondance de blocs a conduit à des améliorations significatives dans ce domaine. Un exemple qui utilise notre approche est donné dans (Linguraru et al. 2008). La mesure de similarité est toujours utilisée (voir eg. Yue et al. 2009). Elle est comparée dans (Zhou et al. 2006) avec des mesures classiques et avec une nouvelle, apprise par un algorithme de boosting. Il ressort de cette comparaison qu'elle est la meilleure parmi les mesures classiques. Soulignons que le modèle statistique adopté se base sur le modèle Rayleigh qui n'est pertinent que si le speckle est pleinement développé et suppose iid les observations des deux bruits η_j^1 et η_j^2 . Un modèle récent qui s'affranchit de ces deux hypothèses est proposé dans (Wachinger et al. 2012). Il s'appuie sur une fdp jointe des deux bruits. Ces derniers sont supposés corrélés et suivent une densité de Nakagami, offrant ainsi plus de flexibilité.

Publications : [25] (cf. page 155), [39].

SEGMENTATION D'IMAGE

Ce chapitre résume mes travaux de recherches sur la problématique de segmentation d'images médicales, travaux effectués au sein du laboratoire Heudiasyc.

5.1 Segmentation automatique des reins

5.1.1 Contexte et Problématique

Le travail décrit dans cette section a été effectué dans le cadre de la thèse de WALA TOUHAMI, en co-direction avec JEAN PIERRE COCQUEREZ. Cette thèse s'inscrit dans le projet de recherche "Système de communication pour la formation en médecine basé sur l'indexation d'image" ; projet en collaboration avec le service radiologique du CHU d'Amiens pour l'étude des *kystes de reins* sur des images tomographiques (CT) de l'abdomen, dans l'objectif de développer ultérieurement un système de recherche d'images par le contenu dédié à l'aide au diagnostic et à la formation des médecins.

Dans ce contexte, la thèse concerne la partie image, plus précisément le problème de la segmentation et de l'identification automatique des régions d'intérêts (les reins sains ou pathologiques en l'occurrence), utiles pour le calcul de la signature de l'index. L'application impose la contrainte du tout automatique des traitements et par conséquent, le problème est difficile. De plus, la grande variabilité anatomique des organes abdominaux augmente sensiblement la difficulté de l'identification automatique des reins. En outre, cette variabilité anatomique est augmentée par la présence de la pathologie ou par l'injection de produits de contraste.

5.1.2 Notre approche

Notre approche pour l'identification automatique des reins s'inscrit dans un cadre statistique. L'observation des images CT de l'abdomen de notre base de données, ainsi que les différentes discussions avec les radiologues nous ont permis de dégager différents descripteurs pertinents caractérisant les reins. En effet, les reins se situent généralement de part et d'autre de la colonne vertébrale et souvent dans la partie inférieure de l'abdomen, sauf dans certains cas où le rein est greffé. Par ailleurs, ils ont en moyenne

une densité de 30 Unités Hounsfield sur les tissus sains, sans kyste ni tumeur. Ces propriétés suggèrent la possibilité de guider leur identification en utilisant un modèle *a priori* sur la position spatiale et un modèle *a priori* sur les niveaux de gris. Nous avons opté pour une solution en deux phases : une première phase de localisation des reins suivie d'une seconde phase de détection. Elles sont précédées d'un prétraitement qui a pour objectif la réduction de la variabilité spatiale de la position des reins dans l'image. Il consiste en la détection de la région abdominale, via des outils de seuillage et de morphologie mathématique. La région détectée sert de support spatial de référence pour tous les autres traitements.

5.1.2.1 Phase de localisation

Nous avons formulé le problème de la localisation des reins comme un problème d'optimisation d'un critère qui permet de positionner deux fenêtres, notées f_g et f_d , respectivement sur les reins gauche et droit. Le critère à optimiser prend en compte des modèles *a priori* de la position spatiale et de la distribution (non-paramétrique) des niveaux de gris des reins sur une image CT d'une part et la propriété de symétrie des reins d'autre part. Cette symétrie sera caractérisée par une mesure de corrélation au sens statistique entre les niveaux de gris des deux fenêtres. Nous proposons comme critère la fonction de coût suivante :

$$J(f_g, f_d, \Theta) = \frac{1}{2} \sum_{i \in \{g, d\}} \{E_{pdf}(f_i) + \lambda_1 E_{spat}(f_i)\} + \lambda_2 E_{sym}(f_g, f_d) , \quad (5.1)$$

où $\lambda_1 \in \mathbb{R}^+$, $\lambda_2 \in \mathbb{R}^+$ sont des hyperparamètres, permettant de pondérer la contribution du terme spatial et du terme de symétrie dans la fonction de coût. Ce critère est minimisé par rapport à la variable $\Theta = \{s_g, s_d\}$ représentant respectivement les centres des fenêtres gauche f_g et droite f_d .

Le premier terme dans (5.1) est un terme d'attache aux données. Il mesure la distance entre la distribution des niveaux de gris du modèle *a priori* et la distribution des niveaux de gris des pixels de la fenêtre courante. Deux mesures sont utilisées, la divergence de Kullback-Leiber et la distance du test de Kuiper. Le second terme est défini comme l'opposé du logarithme de la probabilité *a priori* que la fenêtre courante soit incluse dans le rein¹. Le troisième terme mesure la similarité statistique entre les fenêtres gauche f_g et droite f_d . Nous avons utilisé et comparé l'information mutuelle et le coefficient de corrélation.

5.1.2.2 Phase de détection

Notre démarche consiste à déterminer un ensemble de régions candidates à partir de la fenêtre localisée sur le rein. Ensuite, un ensemble de règles de classification est appris, sur des descripteurs de ces régions, afin de décider lesquelles appartiennent au rein. Le schéma global de cette phase est présenté en figure 5.1. Les étapes principales sont décrites dans les paragraphes suivants :

- **Segmentation** : elle est effectuée, dans un cadre bayésien, en utilisant le critère du maximum a posteriori (MAP) où les observations sont modélisées en utilisant un modèle de mélange de gaussiennes avec un *a priori* modélisé via un

1. D'autres auteurs ont défini des termes équivalents récemment (Shotton et al. 2009, Yang et al. 2012).

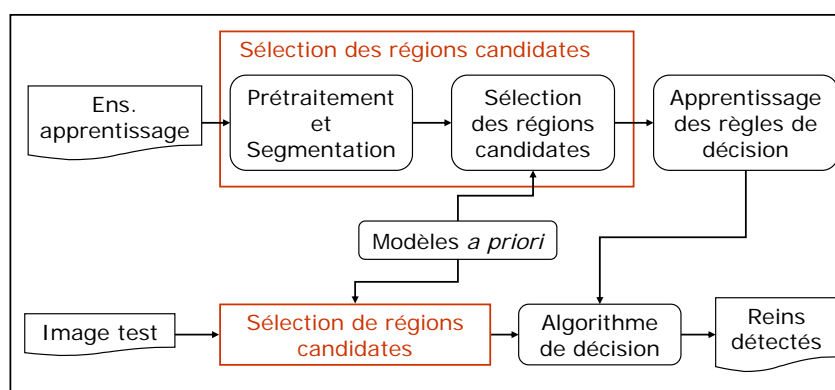


Figure 5.1 – Vue d’ensemble de la phase de détection.

champ de Markov caché. En particulier, nous avons mis en oeuvre l’algorithme HMRF-Expectation-maximization, proposé par [Zhang et al. \(2001\)](#) où l’estimation des paramètres du mélange est obtenue avec l’algorithme EM ([Demps-ter et al. 1977](#)) et la minimisation de la fonction d’énergie est faite à l’aide de l’algorithme déterministe ICM (Iterated Conditional Mode) [Besag \(1986\)](#). Par ailleurs, la minimisation exploite une représentation hiérarchique des données s’appuyant sur la combinaison d’un modèle multi-résolution (pour les observations) et d’un modèle multiéchelle (pour les étiquettes) ([Boukerroui et al. 2003a](#)).

- **Sélection des régions** : ici un ensemble de régions est sélectionné selon un critère de voisinage et de connexité des régions par rapport aux régions ayant une intersection non nulle avec la fenêtre détectée précédemment.
- **Classification** : les régions sélectionnées sont classées comme appartenant au rein ou pas par le biais d’un classifieur binaire. Les régions sont décrites avec un ensemble d’attributs classiques augmenté d’une mesure de voisinage avec la fenêtre localisée et d’une mesure issue de l’a priori spatial. Nous avons utilisé l’arbre de décision binaire CART (Classification and Regression Tree) de [Breiman et al. \(1984\)](#), comme algorithme de classification. Notons que 4 arbres sont appris pour les 4 combinaisons : reins (gauche et droit) et image (avec et sans produit de contraste).

5.1.3 Résultats

Nous avons utilisé une base de données constituée de plus de 500 images, représentatives des différents kystes dont 67 images sont utilisées pour la construction des modèles *a priori* et un ensemble test de 434 images.

Un taux de bonne localisation avoisinant les 97% est obtenu sur les images sans produit de contraste et de 95% pour celles avec un produit de contraste. Aussi la comparaison des différentes combinaisons des mesures dans (5.1), nous permet d’observer que : i) le taux de succès dépend étroitement du paramètre λ_1 , qui pondère la contribution de l’a priori spatial ; ii) le terme de symétrie influe faiblement sur les résultats de l’étape de localisation ; iii) les résultats obtenus avec la mesure de Kuiper sont légèrement meilleurs que ceux obtenus en utilisant la divergence de Kullback-Leiber dans le calcul du terme d’attache aux données.

Pour la phase de détection, des taux entre 90.5% et 94.5% sont obtenus. Nous remarquons une performance légèrement supérieure de notre algorithme sur des images avec produit de contraste. Nous pensons que cette amélioration provient des meilleures performances de l'algorithme de segmentation sur ce type d'images. Quelques exemples d'images de résultats sont montrés en figure 5.2 et d'autre sont disponibles en page 182.

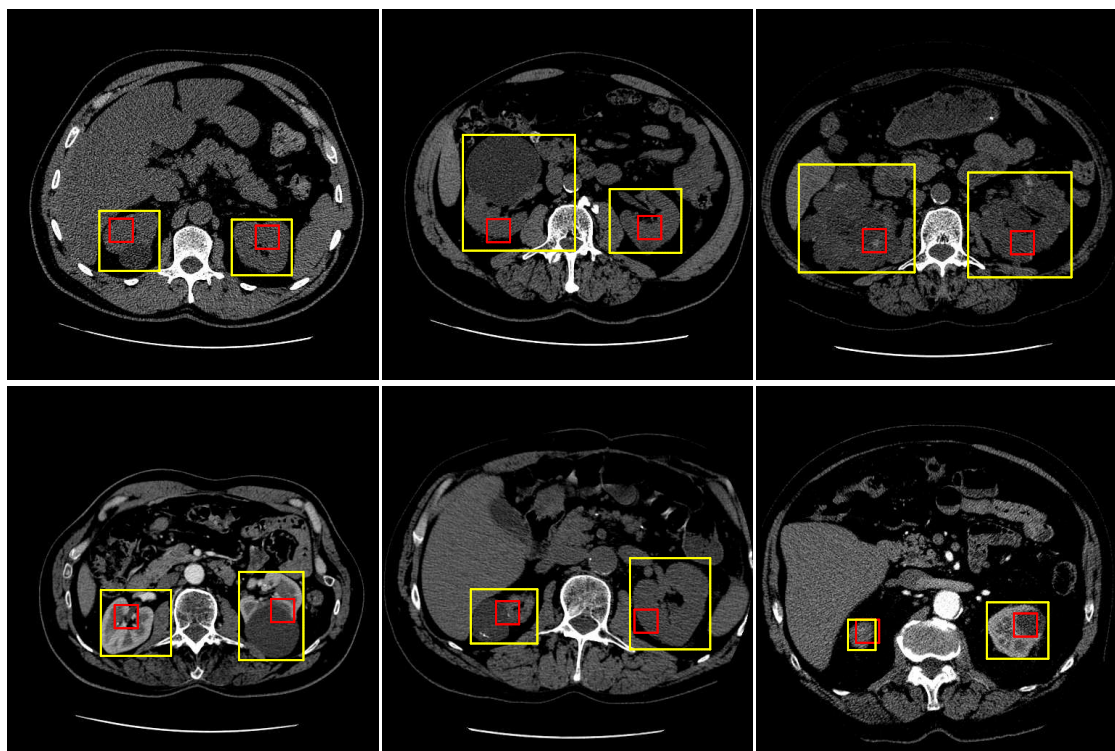


Figure 5.2 – Exemples de détection des reins : images sans (1^{ère} ligne) et avec (2^{ème} ligne) produit de contraste.

5.1.4 Bilan

Les deux contraintes, du tout automatique et de travailler en 2D, ont limité nos choix de méthode. Il serait intéressant de voir le comportement des méthodes récentes de détection d'objets sur ce problème, méthodes utilisées plutôt par la communauté de vision par ordinateur. Notons également que notre approche de détection du rein s'apparente aux méthodes récentes basées sur les super-pixels. Enfin, une formulation variationnelle du problème aurait pu être adoptée offrant plus de flexibilité. Précisons qu'une méthode de classification automatique des reins détectés (pathologique/non pathologique) est proposée dans le cadre de la thèse de WALA TOUHAMI [22].

Publications : [10] (cf. page 169), [24, 23, 33, 22].

5.2 Approches variationnelles en segmentation d'images échographiques

Dans cette section je présente mes travaux de recherches qui traitent de la segmentation d'images échographiques. Ma thèse de doctorat portait également sur cette problématique de recherche. Notons que mes travaux de thèse utilisaient un formalisme bayésien avec un *a priori* markovien; j'ai adopté le cadre variationnel pour mes dernières recherches sur le sujet. Les outils développés intègrent – partiellement – certaines spécificités des données observées.

5.2.1 Introduction

L'imagerie ultrasonore occupe une part importante en imagerie médicale, et constitue la modalité principale dans certains domaines d'application. Elle possède l'avantage d'être non invasive, très pratique et simple d'utilisation. Par ailleurs, la rapidité d'obtention des images (temps réel) fait de l'imagerie ultrasonore une modalité d'image très intéressante pour la visualisation et le diagnostic des tissus en mouvement. Elle est particulièrement bien adaptée à l'étude du cœur (échocardiographie).

En imagerie médicale, l'onde ultrasonore est assimilée à une onde mécanique plane progressive. Lorsque cette onde rencontre une interface séparant deux tissus ayant des caractéristiques acoustiques différentes, une partie de l'onde incidente est réfléchie (échos spéculaires) en direction de la sonde. Parallèlement à ces échos spéculaires, s'ajoutent les ondes rétrodiffusées par la structure microscopique du milieu. La rétrodiffusion est à l'origine du phénomène de speckle, qui caractérise l'imagerie échographique par un aspect granulaire. Le speckle est un *bruit multiplicatif*, fortement corrélé et de *statistique non gaussienne* (voir eg. [Nadarajah 2007](#)). Aussi, comme nous avons eu l'occasion de le rappeler dans le chapitre précédent, cette modalité d'image possède également le désavantage de produire des images très inhomogènes. L'inhomogénéité est essentiellement due au phénomène de l'atténuation des ondes acoustiques en déplacement dans le tissu imagé et à des artefacts tels le renforcement postérieur, l'ombre acoustique ou la réflexion oblique.

Ces caractéristiques s'écartent fortement des hypothèses classiques d'un bruit additif gaussien et blanc supposées dans la plupart des méthodes de traitement, ce qui diminue leur efficacité d'une façon conséquente ([Noble et Boukerroui 2006](#)). En effet, dans le contexte restreint des méthodes de segmentation variationnelles, les approches contour s'appuyant sur le gradient de l'image sont vraisemblablement inadaptées car le RSB (Rapport Signal à Bruit) est très faible. Les modèles régions, offrant plus de flexibilité quant à la modélisation statistique du bruit, sont *a priori* potentiellement meilleurs; ils se heurtent cependant au problème de l'atténuation. En outre, des parties de l'organe imagé sont souvent manquantes à cause des problèmes de réflexion spéculaire, de fente acoustique ou d'atténuation de l'onde acoustique dans la profondeur du milieu observé.

Une solution utilisant l'information de phase, en théorie invariante par rapport au contraste de l'image, est proposée par [Mulet-Parada et Noble \(2000\)](#). Précisons encore une fois que les méthodes de segmentation régions basées sur des modèles statistiques *globaux*, indépendamment du paradigme utilisé, échouent sur ce type de données à cause du problème d'atténuation. Des solutions adaptatives robustes à l'atténuation ont été proposées initialement dans le cadre bayésien classique ([Ashton et Parker 1995](#), [Bou-](#)

kerroui et al. 2003a, 1999). Cela se formalise par l'utilisation des statistiques locales pour l'estimation des paramètres du modèle probabiliste. Ces deux solutions ont fait l'objet d'études pour leur adaptation dans le formalisme des contours actifs implicites dans le formalisme de segmentation variationnelle. Les travaux décrits ci-dessous entrent dans cet axe de recherche.

5.2.2 Approche par phase locale

C'est dans le cadre du projet région ECHOPÉDIA (§ 2.3 page 25), que j'ai dirigé la thèse de M. BELAID, en co-direction avec J-F. LERALLUT, sur la problématique de la segmentation d'images ultrasonore par une approche de phase locale.

Une première idée, intuitive, est de remplacer l'attache aux données des modèles de contours actifs issus du gradient de l'image par un terme reposant sur la phase locale. Parmi ces modèles, on trouve le modèle des Contours Actifs Géodésique (CAG), initialement proposé par Caselles et al. (1997) et largement étudié (voir eg. Aubert et Kornprobst 2006). Un second modèle qui a attiré notre attention est le terme d'alignement proposé dans (Kimmel et Bruckstein 2003) et repris dans (Holtzman-Gazit et al. 2006, Kimmel 2003). En plus de la norme du gradient, ce dernier exploite son orientation. Ils sont rappelés ci-dessous.

5.2.2.1 Contours actifs géodésiques

Le modèle du CAG minimise la fonctionnelle suivante :

$$\mathcal{E}_{GAC}(\mathcal{C}) = \int_0^1 g(|\nabla I(\mathcal{C}(p))|) \left| \frac{\partial \mathcal{C}(p)}{\partial p} \right| dp , \quad (5.2)$$

où \mathcal{C} représente la courbe recherchée, I l'image observée et $g : [0, +\infty[\rightarrow \mathbb{R}^+$ est une fonction strictement décroissante, appelée fonction de détection de contour ou fonction d'arrêt de contour, définie généralement comme (Caselles et al. 1993, Malladi et al. 1995)

$$g = \frac{1}{1 + |\nabla I_\sigma|} \quad \text{ou} \quad g = e^{-|\nabla I_\sigma|} , \quad (5.3)$$

où I_σ est une version filtrée de l'image observée par un filtre gaussien d'écart-type σ . La fonctionnelle définie par (5.2) a l'avantage d'être indépendante de la paramétrisation de la courbe \mathcal{C} et possède de plus une interprétation simple. Elle représente la longueur euclidienne de la courbe pondérée par la fonction g qui contient l'information sur les frontières dans l'image. Le problème est équivalent à trouver une géodésique (courbe de longueur minimale) dans un espace de Riemann défini par une métrique induite par l'image I . Le calcul des variations du problème de minimisation ci-dessus nous conduit à exprimer la solution via l'équation d'évolution suivante :

$$\frac{\partial \mathcal{C}(t)}{\partial t} = g(|\nabla I|) \kappa \vec{n} - \langle \nabla g, \vec{n} \rangle \vec{n} , \quad (5.4)$$

dans laquelle t représente le temps, κ la courbure et \vec{n} le vecteur normal rentrant à la courbe. Le premier terme a pour effet de réduire la longueur de la courbe par un lissage guidé par la courbure κ . Généralement, ce terme peut être augmenté en remplaçant κ par $\kappa + c$, de sorte qu'un rétrécissement ou une expansion de la courbe est possible. Le second terme ne s'exerce que dans un voisinage où $\nabla g \neq 0$. Il a pour rôle d'attirer

puis de caler la courbe sur la frontière réelle de l'objet. La représentation implicite du contour, en utilisant la théorie des ensembles de niveaux, nous permet d'écrire l'équation d'évolution (5.4) sous la forme suivante :

$$\frac{\partial \phi}{\partial t} = \operatorname{div} \left(g(|\nabla I|) \frac{\nabla \phi}{|\nabla \phi|} \right) |\nabla \phi| . \quad (5.5)$$

5.2.2.2 Terme d'alignement

Dans ce modèle, le problème est posé comme étant la recherche d'un contour \mathcal{C} dont la normale \vec{n} est alignée avec un champ de vecteur donné V , d'où le nom de *terme d'alignement*. Dans sa définition la plus générale, le modèle maximise la fonctionnelle donnée par Kimmel et Bruckstein (2003) :

$$\mathcal{E}_A(\mathcal{C}) = \int_0^1 \langle V(\mathcal{C}(p)), \vec{n}(p) \rangle \left| \frac{\partial \mathcal{C}}{\partial p} \right| dp . \quad (5.6)$$

Une version dite robuste est proposée par Kimmel (2003)

$$\mathcal{E}_{AR}(\mathcal{C}) = \int_0^1 |\langle V(\mathcal{C}(p)), \vec{n}(p) \rangle| \left| \frac{\partial \mathcal{C}}{\partial p} \right| dp . \quad (5.7)$$

Les équations d'Euler-Lagrange pour les deux modèles ci-dessus sont respectivement données par :

$$\frac{\delta \mathcal{E}_A(\mathcal{C})}{\delta \mathcal{C}} = \operatorname{div}(V) \vec{n} , \quad (5.8)$$

$$\frac{\delta \mathcal{E}_{AR}(\mathcal{C})}{\delta \mathcal{C}} = \operatorname{sign}(\langle V, \vec{n}(p) \rangle) \operatorname{div}(V) \vec{n} . \quad (5.9)$$

Dans les travaux de Kimmel et co-auteurs, le cas particulier de $V = \nabla I$ est étudié. Ce qui conduit à un résultat très important, vu que la condition de l'extremum donnée en (5.8) nous conduit à $\delta \mathcal{E}_A(\mathcal{C}) / \delta \mathcal{C} = \Delta I \vec{n}$. Ainsi, une interprétation variationnelle du détecteur de contour de Marr et Hildreth (1980) qui définit les contours par le *passage par zéro du Laplacien* est obtenue.

5.2.2.3 Notre proposition

Étant donné que les détecteurs de contours basés sur l'information de phase locale donnent de meilleurs résultats sur les images ultrasonores enveloppes, il est donc naturel de redéfinir la fonction d'arrêt g du modèle de CAG en s'appuyant sur une mesure de phase locale. Nous avons opté encore une fois pour la mesure de Feature Asymmetry donnée en (2.15) dont le principe est détaillée précédemment (§ 2.2.2 page 46). Rappelons que la mesure est normalisée entre 0 et 1. Nous avons testé les deux fonctions suivantes :

$$g = 1 - \text{MFA}^\gamma \quad \text{ou} \quad g = \frac{1}{1 + \gamma \text{MFA}} , \quad (5.10)$$

avec γ un paramètre réel positif. La seconde proposition semble conduire à meilleurs résultats. Dans ce travail, l'extension 2D de la mesure utilise le signal monogénique et nous avons aussi préféré les filtres de Cauchy sur les filtres de log-Gabor pour la partie radiale.

Par ailleurs, le terme d'alignement peut aussi être redéfini en utilisant l'orientation locale estimée par le signal monogénique. Cette dernière est une bonne alternative à une utilisation simple du gradient de l'image, en particulier lorsqu'il s'agit des images ultrasonores. Une façon simple de faire est de définir :

$$V(x) = I_R(x) = (f_e * h_x * I(x), f_e * h_y * I(x))^T, \quad (5.11)$$

où la notation I_R représente la transformée de Riesz de l'image I filtrée par le filtre passe bande 2D isotrope f_e . En fait, cette écriture est très proche de la précédente. En effet, il est facile de vérifier que les deux filtres définissant la transformée de Riesz, h_x et h_y , donnés en (2.7) sont les dérivées partielles du filtre 2D passe-bas isotrope suivant :

$$h(x, y) = \frac{1}{2\pi\sqrt{x^2 + y^2}}. \quad (5.12)$$

Cela nous permet donc d'interpréter la transformée de Riesz comme étant le gradient d'une image lissée. Toutefois, la présence du filtre passe bande f_e dans (5.11) suggère une meilleure estimation, plus robuste à la fois au bruit (hautes fréquences) et à l'atténuation (basses fréquences). Il nous semble que cette proposition est préférable, pour sa simplicité d'interprétation, à l'utilisation du vecteur de phase, défini par Felsberg et Sommer (2004), que nous avons utilisé dans [3].

5.2.2.4 Résultats et Bilan

Nos expérimentations sur des images de synthèse et sur des données réelles ont confirmé la robustesse de notre modèle de segmentation basé sur la phase locale en comparaison au modèle classique du CAG et au modèle région de Sarti et al. (2005), s'appuyant sur une modélisation des statistiques globales par une loi de Rayleigh, augmenté d'un terme de régularisation géodésique. Toutefois, ce travail mérite, ou plutôt nécessite, d'être validé sur une base de donnée plus large. Offrant ainsi une analyse, statistiquement crédible, de l'influence des différents paramètres de contrôle et le choix des filtres passes bandes sur les performances de l'algorithme de segmentation.

Publications : [3] (cf. page 187), [20].

5.2.3 Approches statistiques locales

Dans l'introduction générale de ce document, nous avons évoqué les motivations de l'introduction des modèles de *segmentation région* s'appuyant sur les statistiques locales de l'image. L'introduction de tels modèles dans le formalisme des contours actifs géométriques est un axe de recherche très actif depuis environ cinq années. Dans le paragraphe 1.2.2, une analyse de l'existant et une mise en perspective historique de l'apparition des modèles locaux est donnée. Le bon comportement de ces approches sur des images présentant des dégradés d'intensité suggère leur adéquation à la problématique de segmentation des images ultrasonores.

Nos premiers travaux dans cet axe de recherche reposent sur une modélisation des statistiques locales par une fdp de Rayleigh. Une initiative de mon doctorant A. BELAID a permis une adaptation du modèle global proposé par Sarti et al. (2005) avec les outils proposés par Lankton et Tannenbaum (2008). L'apport des statistiques locales

est clairement mis en évidence [17]. Le modèle, (phase + région) locales, donne les meilleurs résultats.

De mon côté, j'ai travaillé sur la formalisation d'une approche Rayleigh locale, plutôt en m'inspirant du travail de Brox et Cremers (2009) sur le modèle Gaussien local. Ici le calcul des variations de l'énergie à minimiser prend en compte les variations des paramètres du modèle [31]. Une extension de ce travail est apparue dans [15].

Dans ce contexte, notre effort de recherche s'est principalement focalisé pour répondre à ces deux questions :

1. Est-ce qu'un même modèle peut combiner les statistiques locales et globales de l'image ?
2. Comment définir, pour chaque pixel de l'image, une échelle sur laquelle les statistiques locales sont pertinentes pour résoudre le problème de segmentation ?

Ces deux questions constituent la problématique de la thèse de Mlle Q. YANG effectuée sous ma direction. Nos recherches se sont essentiellement concentrées sur la seconde question.

Le problème de sélection d'échelle est très largement étudié par exemple en régression ou en estimation non-paramétrique de densité où le problème à résoudre est souvent formulé comme un compromis biais-variance. En revanche, très peu de travaux existent en image. Parmi les travaux pionniers, on trouve (Boykov et al. 1998, Katkovnik 1999). Katkovnik a beaucoup travaillé sur cet axe de recherche avec comme application type la restauration d'image. Ses recherches sont résumées dans une monographie (Katkovnik et al. 2006)². Parmi les contributions récentes, nous citerons (Azzabou et al. 2007, Duval et al. 2011, Kervrann et Boulanger 2006). Le travail de Katkovnik et al. (2002, 2004) sur le débruitage nous a servi de point de départ.

5.2.3.1 Modèles région

Dans ce qui suit, nous allons rappeler brièvement les termes d'attache aux données des modèles région utilisés. Pour ce faire, les niveaux de gris de l'image observée, I , sont supposés être les réalisations de variables aléatoires indépendantes avec une certaine fdp $p(I(x))$. L'approche adoptée consiste à rechercher la partition du domaine spatiale de l'image, Ω , qui maximise la fonction de vraisemblance des données observées. En adoptant les notations usuelles, cela se ramène à la minimisation d'une fonction d'énergie de forme générale :

$$\mathcal{E}(\phi) = - \sum_{r \in \{i, o\}} \int_{\Omega_r} \log p(I(x)|\mathcal{C}; \theta_r) dx , \quad (5.13)$$

où les deux domaines d'intégration représentent la partition de Ω par la courbe \mathcal{C} en un intérieur Ω_i et un extérieur Ω_e . Ici θ est un vecteur des paramètres du modèle probabiliste. Ainsi le modèle Gaussien de Rousson et Deriche (2002) est obtenu pour :

$$p(I(x)|\mathcal{C}) = \frac{1}{\sqrt{2\pi}\sigma_r} \exp \left\{ -\frac{(I(x) - \mu_r)^2}{2\sigma_r^2} \right\} \text{ avec } \theta_r = (\mu_r, \sigma_r^2)^T , \quad (5.14)$$

2. Voir aussi <http://www.cs.tut.fi/~lasip/>.

et le modèle de Rayleigh de [Sarti et al. \(2005\)](#) pour :

$$p(I(x)|\mathcal{C}) = \frac{I(x)}{\sigma_r^2} \exp \left\{ -\frac{I(x)^2}{2\sigma_r^2} \right\} \text{ avec } \theta_r = \sigma_r^2 . \quad (5.15)$$

Dans les approches dites globales, l'estimation du vecteur des paramètres du modèle est obtenue au sens du maximum de vraisemblance sur tout le domaine Ω_r ([Rousson et Deriche 2002](#), [Sarti et al. 2005](#)). Cela implique l'hypothèse forte que tous les pixels d'un domaine sont identiquement distribués. Cette hypothèse n'est pas réaliste sur des images avec des dégradés d'intensité. Toutefois, elle reste vraie si l'estimation est effectuée localement dans un voisinage, $\mathcal{V}(x)$, centrée autour de chaque pixel du domaine Ω . Cette localité spatiale est spécifiée généralement par un noyau $K(\cdot - x)$ centré autour de la position x . [Brox et Cremers \(2009\)](#) ont proposé le modèle Gaussien local que nous avons utilisé dans [18, 16]. J'ai proposé le modèle de Rayleigh local [31, 15]. Dans mes travaux, un noyau gaussien d'écart type σ_K est utilisé. Cette solution a l'avantage de l'efficacité car elle permet une mise en œuvre récursive ([Geusebroek et al. 2003](#)).

La problématique de recherche telle qu'elle est définie ci-dessus revient donc à trouver un bon critère pour le choix de σ_K en chaque pixel de l'image. J'ai proposé deux critères et une stratégie de sélection de l'échelle en utilisant le principe de l'intersection d'intervalles de confiance (ICI³).

5.2.3.2 Approche LPA-ICI

Dans le cadre de la thèse de Q. YANG, nous avons proposé une solution qui repose essentiellement sur l'approche de débruitage de [Katkovnik et al. \(2002\)](#). Ces derniers ont proposé une méthode de réduction de bruit où l'estimation du niveau de gris en chaque pixel est obtenue par une approximation polynomiale locale (LPA⁴). Sous certaines conditions de régularité de l'image, et pour un modèle de bruit blanc additif, l'analyse de la dépendance de l'Erreur Quadratique Moyenne (EQM) de l'estimateur en fonction de l'échelle du noyau spatial est possible. Cela se formalise comme un problème classique de recherche de compromis variance-biais. Ainsi, une minimisation de l'EQM permet la sélection de l'échelle optimale en chaque pixel de l'image. Néanmoins la solution analytique nécessite la connaissance de l'image sans bruit. [Katkovnik et al. \(2002\)](#) ont proposé une stratégie pour la sélection de l'échelle optimale. Elle repose sur la méthode ICI ([Katkovnik 1999](#)), développée à la suite des travaux pionniers de [Lepskii \(1990\)](#).

La règle ICI recherche la plus grande fenêtre (minimisant donc la variance) où l'approximation polynomiale donne une très bonne estimation du niveau de gris de l'image (minimisant le biais). Pratiquement, soit $\mathbf{h} = \{h_1 < h_2 < \dots < h_J\}$ un ensemble d'échelles, et $\hat{I}_h(x)$ les estimations obtenues par la méthode LPA et $Q_i = Q(h_i)$ les intervalles de confiance correspondant, au niveau $1 - \alpha$, données par ([Katkovnik et al. 2002, 2006](#)) sont :

$$Q(h) = \left[\hat{I}_h(x) - \Gamma \cdot \hat{\sigma}_{\hat{I}_h}(x, h), \hat{I}_h(x) + \Gamma \cdot \hat{\sigma}_{\hat{I}_h}(x, h) \right] , \quad (5.16)$$

où $\hat{\sigma}_{\hat{I}_h}(x, h)$ est une estimation de l'écart type du bruit sur l'estimée et $\Gamma = \gamma + u_{1-\alpha/2}$ avec γ une constante qui dépend de l'ordre de l'approximation LPA et u donne le fractile

3. Pour Intersection of Confidence Intervals.

4. Pour Local Polynomial Approximation.

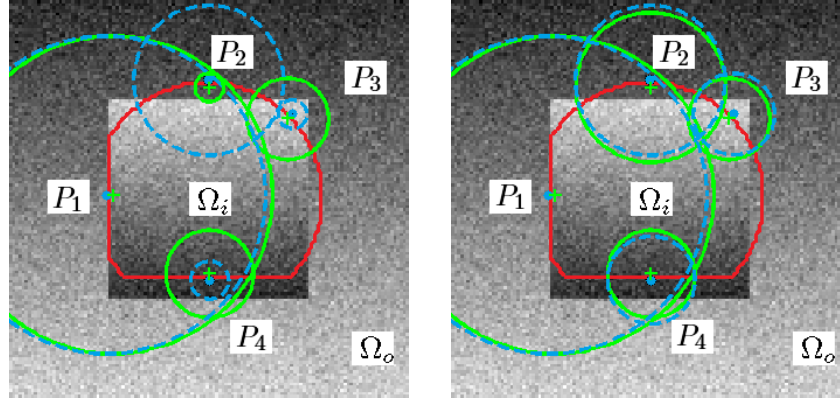


Figure 5.3 – Illustration de la méthode de sélection d'échelle sur une image de synthèse ; quatre paires de points P_1 à P_4 sont montrées. Les points verts $+$ sont dans Ω_i ; les bleus \bullet sont dans Ω_e . Le contour \mathcal{C} est montré en rouge. Les cercles indiquent la taille des noyaux : estimés (à gauche) ; utilisé par l'algorithme de segmentation (à droite).

de la loi normale centrée réduite. Comme la variance est une fonction décroissante de h et le biais est une fonction croissante, il existe nécessairement une échelle pour laquelle l'intersection de deux intervalles de confiance successifs est nulle. Cela est en contradiction avec l'hypothèse que tous les $Q(h_i)$ devraient avoir au moins un point commun qui correspondrait à la vraie valeur, inconnue, de l'estimée. Par conséquent, la méthode ICI définit comme échelle optimale :

$$h^* = \max_{k \in 1 \dots J} \left\{ h_k : \bigcap_{i=1}^k Q(h_i) \neq \emptyset \right\} . \quad (5.17)$$

Mise en œuvre dans le contexte de segmentation : Quelques adaptations sont nécessaires. Notre idée directrice est d'estimer l'échelle optimale conditionnellement à la segmentation courante. Par conséquent l'estimation LPA n'utilise pas tout le voisinage $\mathcal{V}(x)$, mais l'intersection de celui-ci avec la partition courante, i.e. dans le voisinage défini par $\mathcal{V}(x) \cap \Omega_i$ si le pixel appartient à Ω_i ou par $\mathcal{V}(x) \cap \Omega_e$ si le pixel appartient à Ω_e . L'algorithme de segmentation est mis en œuvre par la méthode de la bande étroite. Ainsi une estimation est obtenue pour chaque pixel de la bande. Un prétraitement consistant en un filtrage médian suivi d'un lissage est effectué pour les pixels de chaque bande séparément. Pendant les premières itérations, une seule échelle est utilisée pour l'estimation du modèle paramétrique en chaque pixel le long de la courbe \mathcal{C} . Nous l'avons définie comme étant l'échelle maximale dans la bande étroite le long de la normale à la courbe \mathcal{C} . Les dernières itérations utilisent les échelles estimées pour l'intérieur et pour l'extérieur séparément. Cette stratégie de sélection de l'échelle spatiale est utilisée avec le modèle gaussien de (Brox et Cremers 2009) donné par les équations (5.13) et (5.14) où dans cette dernière le vecteur des paramètres, $\theta_r(x)$ est estimé localement en chaque pixel.

La figure 5.3 montre un exemple d'estimation sur quelques points typiques d'une image de synthèse avec des régions inhomogènes. La figure 5.4 résume les résultats de l'étude expérimentale de l'influence du paramètre de contrôle Γ et du RSB de l'image sur l'estimation de l'échelle locale. On peut noter que h^* est proportionnel à Γ et est inversement proportionnel au RSB et au degré d'inhomogénéité de l'image. Notez par exemple la différence entre les points 2 et 3 ; Contrairement au point 2, la symétrie de l'inhomogénéité au voisinage du point 3 suggère un biais nul pour un voisinage isotrope ; ce qui

induit des échelles plus grandes, réduisant ainsi la variance de l'estimation de la LPA. Quelques exemples de segmentation sont présentés en figure 5.5.

5.2.3.3 Résultats et Bilan

Un des points forts de ce travail est certainement son originalité dans le contexte de la segmentation d'image variationnelle. Piovano et Papadopoulos (2008) ont abordé la problématique de sélection d'échelle dans le contexte qui est le nôtre⁵. La méthode proposée donne des résultats satisfaisants. Elle a cependant le désavantage d'être ad-hoc, ce qui rend difficile l'ajustement du seul paramètre de contrôle. Ce travail nous a servi de méthode de comparaison dans la thèse de QING YANG.

Note approche est testée sur un jeu de données d'images ultrasonore simulées avec le logiciel Field II (Jensen 1996) et sur des images échocardiographiques réelles. Des résultats de comparaison avec quatre algorithmes récents, trois avec échelle fixe (Brox et Cremers (2009), Wang et al. (2009a) et Li et al. (2011a)) et la méthode de Piovano et

5. Travail dont nous avons pris connaissance en fin 2012.

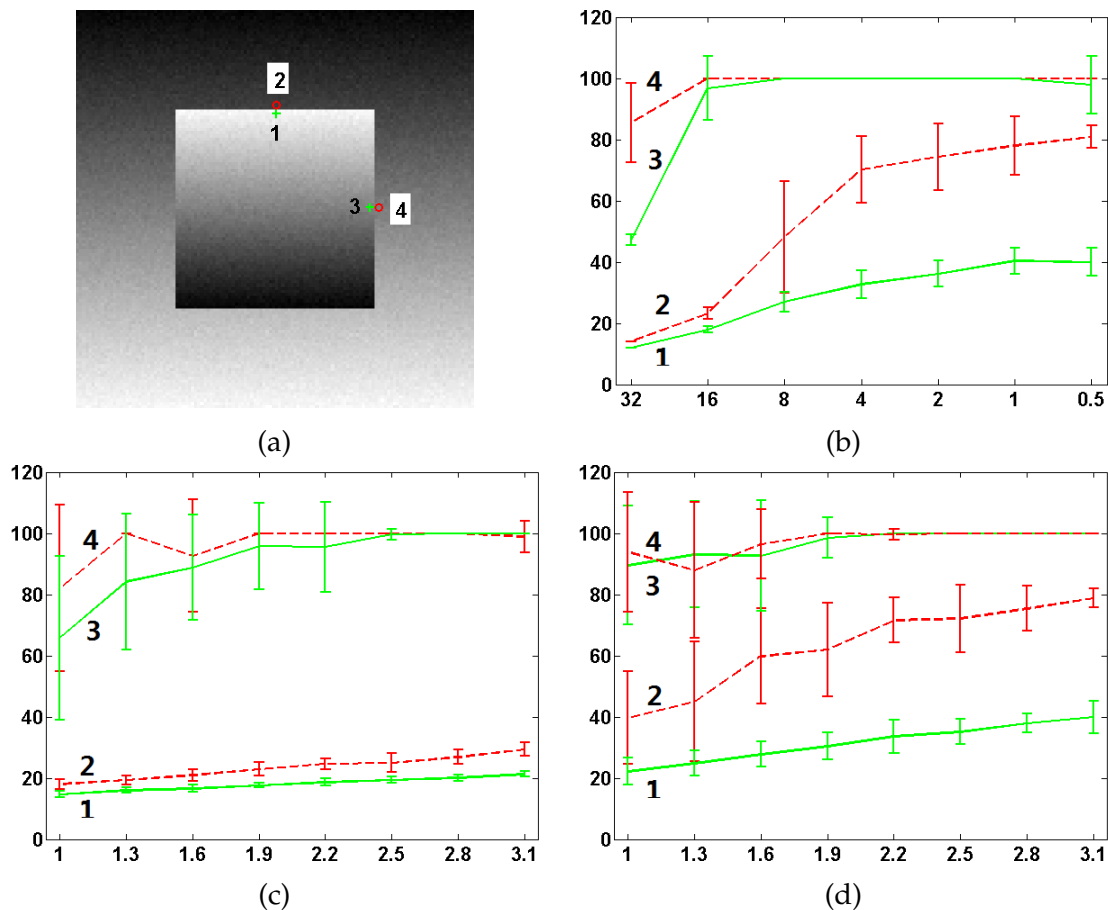


Figure 5.4 – Influence du paramètre de contrôle Γ et du RSB de l'image. Les courbes, de 4 points typiques de l'image (a), montrent l'échelle moyenne estimée en pixel (conditionnellement à la segmentation idéale). La précision de l'estimation est montrée par les barres d'erreur obtenues sur un échantillon de taille 20. (a) image de test de taille 128×128 pour un RSB = 32dB; (b) h^* en fonction du RSB de l'image pour $\Gamma = 2.2$; (c) h^* en fonction de Γ pour un RSB=16dB; (d) h^* en fonction de Γ pour un RSB=4dB.

Papadopoulo (2008) avec une échelle adaptative. L'évaluation quantitative montre un léger avantage de notre approche sur la méthode de Piovano et Papadopoulo (2008). Notons aussi que la méthode récente de Li et al. (2011a), qui permet également l'estimation d'une image de correction du dégradé d'intensité donne des résultats satisfaisants. Soulignons néanmoins que cette dernière échoue à segmenter des images similaires à celles montrées en figure 5.5. En effet, l'hypothèse d'un dégradé lisse sur toute l'image n'est pas vérifiée sur les deux premières images.

La critique majeure de notre approche est que d'une part la méthode est fondée sur l'hypothèse d'un bruit additif gaussien, et d'autre part, la méthode de sélection repose en fait sur un critère de restauration et non pas de segmentation. Une solution est proposée dans la section suivante.

Publications : [16] (cf. page 199), [18].

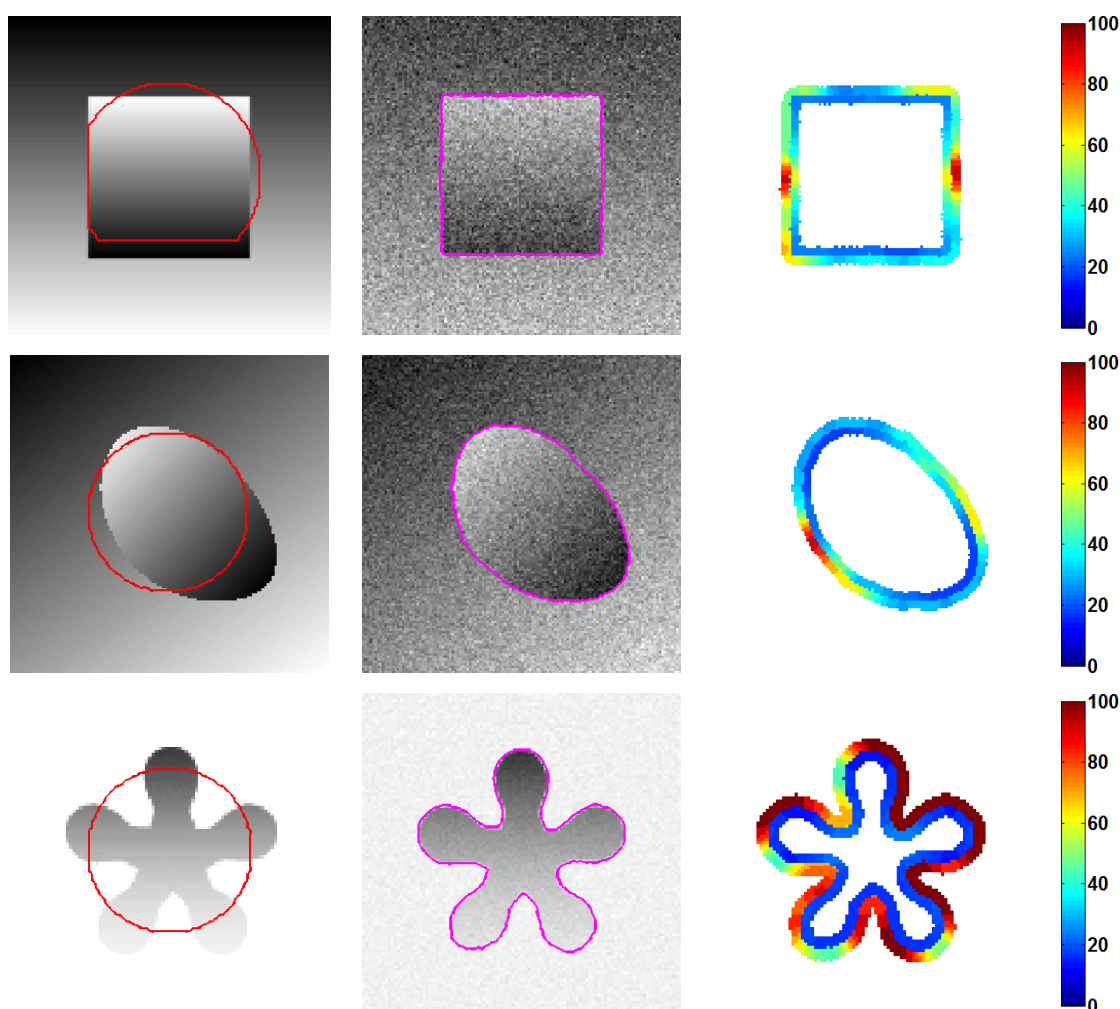


Figure 5.5 – Exemple de segmentation d'images avec des régions inhomogènes en utilisant la méthode LPA-ICI. De gauche à droite : images sans bruit + le contour initial ; résultat de la segmentation sur l'image bruitée ; taille en pixel du noyau spatial estimée le long de la bande étroite du contour final. La taille des trois images est de 128×128 .

5.2.3.4 Approche ICI

Dans le cas général des modèles paramétriques région le vecteur des paramètres θ est généralement estimé par la méthode du maximum de vraisemblance⁶. L'EMV étant un estimateur soit efficace soit asymptotiquement efficace et de plus asymptotiquement gaussien. Il est donc plus naturel que l'estimation de l'échelle locale repose sur une règle d'intersection d'intervalles de confiance sur le paramètre θ . C'est ce que j'ai proposé dans [15]. J'ai tout d'abord appliqué cette stratégie sur le modèle de Rayleigh local pour lequel il existe un estimateur efficace du le paramètre $\theta = \sigma^2$, puis je l'ai appliqué pour le cas gaussien local. Ici, j'ai utilisé les intervalles de confiance sur la moyenne locale, $\mu(x)$, donnée par la fonction pivotale de Student. Notons aussi dans le cas général (i.e non gaussien) l'intervalle de confiance sur la moyenne locale, donné asymptotiquement par une fonction pivotale gaussienne, justifié par le théorème de la limite centrale et le théorème de Slutsky peut être utilisé. Ainsi cette proposition alternative est plus générale ; elle possède le double avantage d'être générique et de rechercher l'échelle locale la plus grande (minimisant la variance de l'estimation des paramètres du modèle de segmentation) où le modèle probabiliste sous-jacent est encore valide (minimisant le biais).

5.2.3.5 Résultats et Bilan

L'approche est testée sur le même jeu de données d'images ultrasonores simulées. J'ai comparé le modèle gaussien local de Brox et Cremers (2009) avec échelle fixe et celui associé avec la stratégie de sélection de Piovano et Papadopoulos (2008) et avec ma proposition (basée sur les intervalles de confiance sur $\mu(x)$). j'ai évalué le modèle de Rayleigh local avec une comparaison au modèle global de Sarti et al. (2005) et le modèle local que j'ai proposé auparavant [31] associé aux deux stratégies de sélection d'échelle. L'évaluation quantitative montre une supériorité du modèle de Rayleigh sur le Gaussien et de ma méthode de sélection sur celle de Piovano et Papadopoulos (2008) et ce pour les deux modèles (voir table des résultats page 214). Ce travail ouvre de nouvelles perspectives, que j'estime très prometteuses.

Publications : [15] (cf. page 205), [31].

6. Les estimateurs issus de la méthode des moments sont aussi asymptotiquement gaussien.

SCHÉMAS NUMÉRIQUES POUR LE GVF

Dans les premiers modèles de contours actifs, l'énergie externe utilise uniquement l'information du gradient de l'image I et est définie comme $-f(\cdot)$, où $f(\cdot)$ est une carte de contours de l'image $I(\cdot)$, habituellement définie comme étant la norme du gradient d'une version filtrée de l'image originale par un filtre gaussien d'écart type σ ; c'est-à-dire $|\nabla G_\sigma(\cdot) * I(\cdot)|^2$. La valeur de σ permet le contrôle de la zone d'influence du contour, dite également zone d'attractivité. Les vecteurs gradients de la carte de contour pointent vers les frontières et sont normaux aux points de discontinuités de l'image. Le champ de vecteurs ∇f est le champ de force externe responsable de l'attraction du modèle déformable vers les discontinuités désirées. Cependant, cette force est importante seulement au voisinage immédiat de la frontière. De tels modèles rencontrent de nombreuses difficultés, parmi lesquelles leur sensibilité à l'initialisation, en raison de l'étroite zone d'influence de la force externe, et surtout leur difficulté à converger vers les concavités des frontières. Cela impose donc une initialisation du contour actif très proche de la solution.

Plusieurs chercheurs ont souligné ces limitations et ont proposé des solutions alternatives afin d'accroître la zone d'influence du modèle et de réduire ainsi sa sensibilité à l'initialisation (voir par ex. [Cohen et Cohen 1993](#), [Cohen 1991](#), [Leroy et al. 1996](#)). De récents travaux proposent également des solutions au problème de convergence vers des frontières concaves. Nous citerons, par exemple, le modèle des particules chargées (Charged-Particle Model) ([Jalba et al. 2004](#)), le modèle de contour actif basé sur les forces de charges électrostatiques (CACE) ([Yang et al. 2006](#)) et le modèle de convolution avec un champ de vecteur (Vector Field Convolution) ([Li et Acton 2007](#)). Dans ce dernier, la force externe est obtenue via une simple convolution de la carte de contours avec un noyau de convolution vectoriel défini par l'utilisateur. Une alternative intéressante est le modèle fondé sur les forces magnétostatiques ([Xie 2010](#), [Xie et Mirmehdi 2008](#)).

La première et la plus populaire des solutions proposées dans la littérature est toutefois le champ de flot de vecteur gradient ou 'Gradient Vector Flow' (GVF) ([Xu et Prince 1998b](#)) et sa généralisation, le GGVF ([Xu et Prince 1998a](#))¹. Depuis sa publication il y a une quinzaine d'années, le GVF a été utilisé et adapté à différents modèles et problématiques relevant du traitement d'images. Nous citerons à titre d'exemples, la segmentation ([Paragios 2002](#), [Paragios et al. 2004](#), [Tauber et al. 2005b](#), [Xu et Prince 1998b](#)), le suivi ([Ray et Acton 2004](#)), la segmentation en utilisant des relations spatiales floues (Col-

liot et al. 2006), le recalage (Guo et Lu 2006), la détection des points saillants (Donner et al. 2007, Stöttinger et al. 2008) et la squelettisation (Hassouna et Farag 2007).

Un certain nombre de variantes et des extensions ont été proposées dans la littérature. Par exemple, le GVF directionnel (G & DVF) renforce les capacités de convergence vers les frontières désirées en ajoutant des informations directionnelles *a priori* (Zhu et al. 2010). Le modèle GVF directionnel dynamique (DDGVF) est une variante intéressante qui utilise également des informations directionnelles. Il est conçu pour être attiré vers les frontières intérieures ou extérieures exclusivement, utile pour la segmentation des objets ayant une forme complexe (Cheng et Foo 2006). Plus récemment, une proposition tenant en compte de la structure locale de l'image dans le processus de diffusion est proposée dans (Wang et al. 2010). Par conséquent, la diffusion est anisotrope (constante pour le GVF et non linéaire pour la GGVF). Le champ de force obtenu améliore la convergence vers de très longues concavités et préserve les bords faibles. Le Normal-GVF restreint la diffusion à la direction normale de l'arête (Ning et al. 2007). Ce dernier est à l'origine de la proposition de Guillot et Le Guyader (2009) d'une variante originale en réduisant le nombre d'inconnues, en supposant que le champ de vecteurs recherché est le gradient d'une fonction scalaire, et la diffusion est dans la direction normale de l'arête. Enfin des méthodes d'initialisation automatique du modèle de contour actif du type GVF peuvent être trouvées dans (Li et al. 2005, Tauber et al. 2005a,b) et la normalisation du champ de vecteurs obtenu est analysée dans (Qin et al. 2013).

La popularité du GVF est certainement due à son efficacité. Le principal inconvénient de l'approche et de sa généralisation est leur coût en temps de calcul, et sa conséquence sur la zone d'attractivité. En effet, traditionnellement, le GVF est obtenu via la minimisation itérative d'une fonctionnelle. Ainsi, la zone d'influence du flot résultant dépend bien sûr du nombre d'itérations utilisé. Très récemment, les techniques de résolution multigrille ont été utilisées avec succès pour résoudre le problème du GVF et ses généralisations (Han et al. 2007, Vidholm et al. 2006).

6.1 Formulation du GVF

L'approche du flot de vecteur gradient permet l'extension de la zone d'influence de la carte de contour en utilisant un processus de diffusion. Soit $f : \Omega \rightarrow \mathbb{R}$ une carte de contour de support spatiale $\Omega \subset \mathbb{R}^d$. Le flot est défini comme le champ de vecteurs $v(x) : \Omega \rightarrow \mathbb{R}^d$ qui minimise la fonctionnelle suivante (Xu et Prince 1998b) :

$$\mathcal{E}(v) = \int_{\Omega} g(x) |\nabla v|^2 + h(x) |v - \nabla f|^2 dx, \quad (6.1)$$

où $g(x)$ et $h(x)$ sont des fonctions à valeurs positives. Selon le choix des deux fonctions de pondérations $g(x)$ et $h(x)$, Xu et Prince (1998a,b) ont proposé deux champs de flot de vecteur gradient :

1. **Le flot de vecteur gradient (GVF)** : il est défini pour le cas où $g(x) = \mu$ et $h(x) = |\nabla f|^2$. μ est une constante positive qui permet le contrôle du terme de régularisation dans (6.1).
2. **Le flot de vecteur gradient généralisé (GGVF)** : afin d'éviter un lissage excessif dans le voisinage des discontinuités détectées, et de renforcer la régularisation

1. Voir aussi le site web sur le GVF <http://www.iac.ece.jhu.edu/static/gvf/>.

dans les régions lointaines, le choix suivant est proposé par [Xu et Prince \(1998a\)](#) :

$$g(x) = \exp \left\{ -\frac{|\nabla f|}{\kappa} \right\} , \quad h(x) = 1 - g(|\nabla f|) , \quad (6.2)$$

avec κ un paramètre positif.

L'idée de base de la formulation variationnelle (6.1) est d'imposer une certaine régularité à la solution dans les régions lointaines des données (premier terme) et d'être très proche du champ initial là où des mesures fiables sont disponibles (second terme). Le calcul des variations du problème de minimisation ci-dessus nous conduit à exprimer la solution via l'équation d'évolution suivante ([Xu et Prince 1998a,b](#)) :

$$\frac{\partial}{\partial t} v(x, t) = \nabla \cdot (g(x) \nabla v) - h(x)(v - \nabla f) , \quad (6.3)$$

avec des conditions aux limites de type Neumann, et $v(x, 0) = \nabla f$ comme condition initiale. [Xu et Prince \(1998b\)](#) ont proposé un schéma de résolution numérique explicite. Les schémas explicites sont très populaires pour la résolution des équations aux dérivées partielles, principalement en raison de leur simplicité. En contre partie à cette simplicité, la condition de stabilité impose des pas en temps très petits (en 2D, $\tau < 1/4$), ce qui conduit à des coûts de calcul importants.

6.2 Contributions

J'ai commencé à m'intéresser à la méthode du GVF suite au stage de Master de Rachid BENMOKHTAR. Je me suis rendu compte qu'il y avait une erreur dans la condition de stabilité du schéma explicite donné par [Xu et Prince \(1998b\)](#). Dans ce travail, j'ai proposé et comparé l'efficacité d'autres schémas numériques alternatifs, qui tous permettent l'utilisation d'un plus grand pas en temps :

1. Semi-implicite avec une résolution par la Transformée en Cosinus discrète (GVF seulement).
2. Directions alternées explicite ou ADE pour 'Alternating Direction Explicit'. Une variante ADES est proposée.
3. Méthodes d'approximation par factorisation du schéma implicite :
 - (a) Additive AOS (pour 'Additive Operator Splitting');
 - (b) Multiplicative LOD (pour 'Locally One Dimensional');
 - (c) Combinée AMOS (pour 'Additif Multiplicative Operator Splitting');
 - (d) AFI (pour 'Approximation Factorisation Implicit') variante de LOD.

Tous ces schémas sont du second ordre en espace et sont du premier ordre en temps, et sont inconditionnellement stable (à l'exception du semi-implicite). Toutes les méthodes ont une complexité $O(N^2)$ à l'exception du schéma semi-implicite utilisant la DCT où la complexité est de l'ordre $O(N^2 \log N)$ pour une image de taille $N \times N$. Le temps de calcul pour une itération du schéma ADE est approximativement le même que pour le schéma explicite, et il est de 2 à 3 fois plus important pour les schémas ADES, LOD, AOS et AFI. La méthode AMOS a besoin en revanche d'un effort 5 fois plus important. Ces résultats sont basés sur une moyenne de 600 itérations et pour 8

dimensions différentes de l'image. Cela impose donc un minimum de pas en temps entre 0,25 et 1,5 pour que les alternatives proposées soient plus rapides que l'explicite. Il est à retenir cependant qu'un pas en temps très grand introduit une perte en précision, et dans certaines approximations une perte d'isotropie.

Le schéma utilisant la DCT ayant une complexité plus importante et une condition de stabilité plus restrictive, généralement un pas en temps max de 3 à 5, le gain en temps de calcul est moindre (de 2 à 3 par rapport au schéma explicite). Il convient de souligner toutefois que cette méthode peut être très intéressante dans la mesure où une approximation d'ordre supérieur est nécessaire pour la discrétisation du Laplacien.

Une seconde contribution est la mise au point d'une stratégie de comparaison quantitative dédiée. En effet, lors de la mise au point d'une méthode de quantification de la précision des différents schémas numériques, nous devons également garder à l'esprit notre objectif final quant à l'utilisation du GVF. En effet, généralement, dans la plupart de ses utilisations, l'information la plus importante est l'orientation des vecteurs du flot. L'amplitude est rarement utilisée dans des applications de segmentation par exemple. Pour ce faire, j'ai synthétisé une carte de contours pour laquelle une vérité terrain sur l'orientation théorique du champ de vecteurs en chaque point de l'image est disponible. Deux critères sont alors utilisés pour la comparaison des différentes méthodes : l'erreur angulaire absolue moyenne estimée du champ de vecteurs obtenu par rapport à la vérité terrain, et la valeur de la fonction d'énergie \mathcal{E} donnée par l'équation (6.1).

Mes résultats montrent que les schémas AFI et ADES sont très satisfaisants et sont meilleurs, en comparaison aux autres propositions. Comme prévu, le schéma AMOS est en général meilleur que les deux qui le composent (LOD et AOS). La précision de la méthode ADES est acceptable, même pour des pas en temps grands, et ce pour toutes les combinaisons des paramètres. En effet, l'erreur sur l'orientation des vecteurs gradients est du même ordre que la méthode explicite pour $t > 50$ avec $\tau = 10$, et pour $t > 100$ pour $\tau = 25$ (autrement dit seulement au bout de 4 à 5 itérations).

J'ai également implémenté le solveur multigrille, proposé par [Han et al. \(2007\)](#) pour résoudre le problème de GGVF. L'application du FMG-GS (1,2,2) génère une erreur angulaire de $1,9^\circ$ pour $\kappa = 0.2$. L'erreur de la méthode ADES est inférieure à $1,9^\circ$ après quatre itérations avec $\tau = 15$. Un facteur d'accélération de l'ordre de 60 est donc obtenu en comparaison par rapport à la méthode explicite. Enfin, notons que depuis la publication de mon travail, une méthode de résolution du type lagrangien augmenté est proposée pour la résolution du GVF ([Li et al. 2011b](#)). Ce travail utilise une résolution des sous problèmes par la FFT. Ainsi, il n'est pas adapté pour la résolution du problème du GGVF.

Publications : [2] (cf. page 219), [32, 21].

BILAN & PERSPECTIVES

Ce document, composé de deux parties, présente une synthèse de mon activité professionnelle depuis la soutenance de ma thèse. La première partie revient sur mon parcours, détaillant mes implications dans les divers enseignements et comités de relecture ainsi que les projets de recherche et les directions de thèses auxquels j'ai pris part.

La seconde partie présente mes principales contributions dans une chaîne de traitement qui a pour objectif de résoudre un problème de segmentation d'image. Les applications visées sont toujours de nature médicale et par conséquent ma recherche est orientée, contrainte par la nature des données et motivée par l'application.

Mes contributions sont essentiellement algorithmiques et sont, pour la plupart d'entre elles, appliquées sur des images ultrasonores. D'autres répondent à un problème méthodologique spécifique de traitement d'image, comme mon travail sur le "Gradient Vector Flow". À noter également mes contributions à des travaux de synthèse, sur les filtres en quadrature et sur la segmentation d'image ultrasonore. Dans tout le document, et dans la mesure du possible, j'ai essayé de décrire mes contributions passées dans le contexte de l'état de la recherche d'aujourd'hui. J'ai aussi fourni des éléments d'analyse sur les travaux présentés.

Il apparaît notamment que mes récents travaux sur la segmentation d'image ultrasonore nécessitent certainement plus de validation, à l'image de beaucoup de travaux de traitement d'images médicales. Il est certain que la communauté de traitement d'images médicales, et bio-médicales en général, a beaucoup de retard sur le plan de la validation et de la comparaison des travaux de recherche sur des bases d'images représentatives de la problématique à résoudre. En effet, ce n'est qu'à partir de 2006 que des challenges ont été organisés, dans le cadre des conférences MICCAI puis ISBI, sur une problématique donnée¹. Cette prise de conscience est survenue après celle des communautés de la vision par ordinateur et de l'apprentissage statistique.

Cette synthèse me permet de mettre en évidence certaines perspectives, dans la continuité des travaux présentés, qu'il me paraît souhaitable d'explorer à court terme. Elles sont énoncées ci-dessous en soulignant également l'importance de l'évolution des

1. <http://www.grand-challenge.org/>

approches alternatives.

Le travail de Sommer et co-auteurs sur le signal monogénique a ouvert un axe de recherche très actif sur l'utilisation de la transformée de Riesz, et sur les ondelettes complexes en général en traitement d'image avec des applications en imagerie médicales ou couleur (données vectorielles). Cependant, très peu de travaux existent encore sur l'influence de la partie radiale des filtres et sur les mesures de détection. Notons toutefois le travail de [Mellor et al. \(2008\)](#) sur une proposition d'une famille de filtres quasi-invariants en échelle. Une contribution très intéressante, bien que ancienne, est le travail de [Marroquin et al. \(1997\)](#) sur les filtres en quadrature robustes. Quant aux mesures de détection, elles sont toutes des variantes des mesures de Kovess présentées au chapitre 2. On peut s'interroger néanmoins sur l'efficacité de tels outils sur une classe d'images plus importante. En effet, les performances de détecteurs de contour plus sophistiqués, par patch, qui intègrent différentes mesures (texture, couleur, etc) pour apprendre un modèle de frontières dédié à une classe d'images sont largement meilleurs ([Arbelaez et al. 2011](#), [Martin et al. 2004a](#)). Ces travaux montrent notamment que la texture mérite une attention particulière pour obtenir de bon résultats ([Martin et al. 2004a](#)). Notons toutefois que de tels modèles intègrent aussi des mesures issues des filtres en quadrature. Une extension à des données statio-temporelles d'un modèle simplifié est formulée dans ([Stein et Hebert 2006](#)).

Mon travail sur la sélection d'échelle spatiale locale ouvre de nombreuses perspectives. La première est d'ordre pratique vu que pour le moment je ne me suis pas préoccupé de l'efficacité de l'implémentation de cette approche. Son extension aux critères de segmentation intégrant des modèles de bruit non paramétriques est également à envisager. Soulignons également que l'utilisation conjointe des statistiques locales, à une échelle fixe, et globales en segmentation variationnelle d'image est très peu explorée ([Bresson et Chan 2008](#), [Wang et al. 2009b](#)). Le problème du compromis entre les deux termes, global et local, et celui du choix de la taille optimale de la fenêtre locale ou d'une fenêtre adaptative méritent d'être étudiés. Toutefois, au-delà de toute considération de complexité, on peut se demander si les modèles de segmentation fondés sur des mesures mono-pixel ne seraient pas surpassés par la récente proposition d'un modèle de segmentation s'appuyant sur des comparaisons de blocs ([Jung et al. 2012](#)). À l'instar de l'impact du travail de [Buades et al. \(2005b\)](#) sur le débruitage d'image, je prévois une influence similaire sur la segmentation d'image.

Ce document étant guidé par mes contributions, je n'ai pas mentionné certains travaux sur la segmentation d'image, ou plus précisément, des axes de recherche, que je pense être importants. En effet, j'ai très peu parlé des méthodes par graphes alors qu'elles ont été très présentes lors de la dernière décennie. Elles sont autant adaptées pour de la segmentation interactive que les modèles déformables². L'apparition récente des méthodes dites "Superpixels" ouvre de nouvelles perspectives (voir eg. [Achanta et al. 2012](#)). Ces approches sont particulièrement attractives sur des images de grandes tailles. L'utilisation de cette nouvelle représentation de l'image en segmentation, ou pour résoudre d'autres problématiques image, est un axe de recherche porteur. Récemment, la problématique de la segmentation d'image a été souvent associée, en vision par ordinateur, à celle de la détection d'objets ([Shotton et al. 2009](#), [Yang et al. 2012](#)). Aussi, le terme de segmentation sémantique est apparu, qui induit une classification des régions segmentées.

2. Toutefois des approches alternatives plus rapides existent (voir eg. [Criminisi et al. 2008](#)).

Mon souhait et mon ambition personnelle ont toujours été de faire de la recherche pour améliorer notre quotidien. Il apparaît clairement au travers de mon parcours professionnel et de cette synthèses scientifique, que mes contributions les plus influentes sont celles qui portent sur des traitements des images échographiques et plus précisément sur la problématique de la segmentation. Mon expertise dans ce domaine est reconnue au niveau international. En effet, un grand pourcentage de mes sollicitations comme expert ou rapporteur concerne cet axe de recherche. Par conséquent, mon premier souhait est de poursuivre ma recherche dans cette thématique. L'activité "imagerie médicale" étant arrêtée au sein du laboratoire Heudiasyc (depuis maintenant deux années), cette perspective est donc conditionnée par un changement de laboratoire, intéressé par mon expertise pour animer une activité de recherche en imagerie ultrasonore. Cela dépendra des opportunités à venir et je ne limite pas ma recherche au seul territoire français.

En revanche, poursuivre ma recherche à Heudiasyc nécessite une mobilité thématique, pour me rapprocher de l'activité vision par ordinateur et de l'apprentissage statistique de mon laboratoire. J'ai déjà entamé une collaboration avec VINCENT FRÉMONT, Maître de conférence, HDR, à Heudiasyc dans le cadre d'un contrat de recherche avec le CETIM sur la "perception multicateurs pour la détection de risques de collision". La problématique de l'application s'apprête à l'investigation de certaines perspectives énoncées ci-dessus, plus précisément à la segmentation sémantique d'images et à l'utilisation des modélisations Superpixels.



RÉFÉRENCES

- R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, et S. Ssstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(11) :2274–2282, 2012.
- A. Adam, R. Kimmel, et E. Rivlin. On scene segmentation and histograms-based curve evolution. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(9) :1708–1714, 2009.
- M. Alessandrini, A. Basarab, H. Liebgott, et O. Bernard. Myocardial motion estimation from medical images using the monogenic signal. *IEEE Trans. Image Process.*, 22(3) :1084–1095, 2013a.
- M. Alessandrini, O. Bernard, A. Basarab, et H. Liebgott. Multiscale optical flow computation from the monogenic signal. *IRBM*, 34(1) :33 – 37, 2013b.
- L. Alvarez, F. Guichard, P.-L. Lions, et J.-M. Morel. Axioms and fundamental equations of image processing. *Archive for Rational Mechanics and Analysis*, 123 :199–257, 1993.
- S. Ando. Consistent gradient operators. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(3) :252–265, 2000.
- E. Angelini, Y. Jin, et A. Laine. State of the art of level set methods in segmentation and registration of medical imaging modalities. Dans J. Suri, D. Wilson, et S. Laxminarayan, diteurs, *Handbook of Biomedical Image Analysis*, Topics in Biomedical Engineering International Book Series, pages 47–101. Springer US, 2005. ISBN 978-0-306-48607-4.
- S. Angenent, E. Pichon, et A. Tannenbaum. Mathematical methods in medical image processing. *Bulletin of the American Mathematical Society*, 43(3) :365–396, 2006.
- P. Arbelaez, M. Maire, C. Fowlkes, et J. Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5) :898–916, Mai 2011.
- E. Ashton et K. Parker. Multiple resolution Bayesian segmentation of ultrasound images. *Ultrasonic Imag.*, 17(2) :291–304, 1995.
- G. Aubert, M. Barlaud, O. D. Faugeras, et S. Jehan-Besson. Image segmentation using active contours : Calculus of variations or shape gradients? *SIAM J. Applied Mathematics*, 63(6) : 2128–2154, 2003.
- G. Aubert et P. Kornprobst. *Mathematical Problems in Image Processing : Partial Differential Equations and the Calculus of Variations*, volume 147. Springer, Applied Mathematical Sciences, second edition dition, 2006.

- J.-F. Aujol, G. Aubert, et L. Blanc-Féraud. Wavelet-based level set evolution for classification of textured images. *IEEE Trans. Image Process.*, 12(12) :1634–1641, 2003.
- I. B. Ayed, N. Hennane, et A. Mitiche. Unsupervised variational image segmentation/classification using a weibull observation model. *IEEE Trans. Image Process.*, 15(11) : 3431–3439, 2006.
- I. B. Ayed, A. Mitiche, et Z. Belhadj. Multiregion level-set partitioning of synthetic aperture radar images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(5) :793–800, 2005.
- N. Azzabou, N. Paragios, F. Guichard, et F. Cao. Variable bandwidth image denoising using image-based noise models. Dans *CVPR*, pages 1–7. IEEE Computer Society, 2007.
- E. Bae, J. Yuan, et X.-C. Tai. Global minimization for continuous multiphase partitioning problems using a dual approach. *Int. J. Comput. Vis.*, 92(1) :112–129, 2011.
- S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, et R. Szeliski. A database and evaluation methodology for optical flow. *Int. J. Comput. Vis.*, 92(1) :1–31, 2011.
- S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, et R. Szeliski. A database and evaluation methodology for optical flow. Dans *IEEE Int. Conf. Comput. Vis.*, pages 1–8, 2007.
- A. Belaid, D. Boukerroui, Y. Maingourd, et J.-F. Lerallut. Phase-based level set segmentation of ultrasound images. *IEEE Trans. Inf. Technol. Biomed.*, 15(1) :138–147, 2011.
- J. Besag. On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society. Series B*, 48(3) :259–302, 1986.
- A. Bharath et J. Ng. A steerable complex wavelet construction and its application to image denoising. *IEEE Trans. Image Process.*, 14(7) :948–959, 2005.
- M. Black et A. Jepson. Estimating optical flow in segmented images using variable-order parametric models with local deformations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(10) :972–986, 1996.
- A. Blake et M. Isard. *Active Contours*. New York : Springer-Verlag, 1998.
- E. Borenstein et S. Ullman. Combined top-down/bottom-up segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(12) :2109–2125, 2008.
- D. Boukerroui. A local Rayleigh model with spatial scale selection for ultrasound image segmentation. Dans R. Bowden, J. Collomosse, et K. Mikolajczyk, éditeurs, *British Mach. Vis. Conf.*, pages 84–84, Surrey, UK, Septembre 2012.
- D. Boukerroui, A. Baskurt, A. Noble, et O. Basset. Segmentation of ultrasound images—multiresolution 2D and 3D algorithm based on global and local statistics. *Pattern Recognit. Lett.*, 24 :779–790, 2003a.
- D. Boukerroui, O. Basset, A. Baskurt, et A. Noble. Segmentation of echocardiographic data. Multiresolution 2D and 3D algorithm based on gray level statistics. Dans *MICCAI*, pages 516–524, Cambridge, UK, 1999. Springer-Verlag.
- D. Boukerroui, O. Basset, N. Guérin, et A. Baskurt. Multiresolution texture based adaptive clustering algorithm for breast lesion segmentation. *European J. Ultrasound*, 8(2) :135–144, 1998.
- D. Boukerroui, J. Noble, et M. Brady. Velocity estimation in ultrasound images : a block-matching approach. Dans *Info. Proc. Med. Imag.*, volume 2732 de *Lect. Note Comput. Sci*, pages 586–598, Ambleside, UK, 2003b.

- D. Boukerroui, J. A. Noble, et M. Brady. *Frontiers in robotics research. Chapitre On the Selection of Band-Pass Quadrature Filters*, pages 67–111. Nova Publishers, Max A. Denket édition, 2006.
- D. Boukerroui, J. A. Noble, M. C. Robini, et M. Brady. Enhancement of contrast regions in suboptimal ultrasound images with application to echocardiography. *Ultrasound Med. Biol.*, 27 (12) :1583–1594, Décembre 2001.
- E. Bourennane, P. Gouton, M. Paindavoine, et F. Truchetet. Generalization of Canny—Deriche filter for detection of noisy exponential edge. *Signal Process.*, 82(10) :1317–1328, Octobre 2002.
- Y. Boykov et G. Funka-Lea. Graph cuts and efficient ND image segmentation. *Int. J. Comput. Vis.*, 70(2) :109–131, 2006.
- Y. Boykov et M. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. Dans *IEEE Int. Conf. Comput. Vis.*, pages 105–112, Vancouver, BC, Canada, 2001.
- Y. Boykov et V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. Dans *IEEE Int. Conf. Comput. Vis.*, pages 26–33, 2003.
- Y. Boykov, V. Kolmogorov, D. Cremers, et A. Delong. *An Integral Solution to Surface Evolution PDEs via Geo-Cuts*, volume 3953 de *Lecture Notes in Computer Science*. Springer, Juin 2006.
- Y. Boykov, O. Veksler, et R. Zabih. A variable window approach to early vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(12) :1283–1294, 1998.
- L. Breiman, J. H. Friedman, R. A. Olshen, et C. J. Stone. *Classification and Regression Trees*. Wadsworth, 1984. ISBN 0-534-98053-8.
- X. Bresson et T. F. Chan. Non-local unsupervised variational image segmentation models. Rapport Technique 08-67, UCLA C.A.M. Report, 2008.
- X. Bresson, S. Esedoglu, P. Vanderghenst, J. Thiran, et S. Osher. Fast global minimization of the active contour/snake model. *J. Math. Imag. Vis.*, 28(2) :151–167, 2007.
- T. Brox et D. Cremers. On the statistical interpretation of the piecewise smooth Mumford-Shah functional. Dans *Int. Conf. on Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 4485 de *Lecture Notes in Computer Science*, pages 203–213. Springer, Mai 2007.
- T. Brox et D. Cremers. On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional. *Int. J. Comput. Vis.*, 84(2) :184–193, 2009.
- A. Buades, B. Coll, et J.-M. Morel. A non-local algorithm for image denoising. Dans *Proc. of IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, volume 2, pages 60–65, 2005a.
- A. Buades, B. Coll, et J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation*, 4(2) :490–530, 2005b.
- M. Cabezas, A. Oliver, X. Llado, J. Freixenet, et M. B. Cuadra. A review of atlas-based segmentation for magnetic resonance brain images. *Computer Methods and Programs in Biomedicine*, 104 (3) :e158 – e177, 2011.
- J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8 (6) :679–698, 1986.
- V. Caselles, F. Catté, T. Coll, et F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1) :1–31, 1993.
- V. Caselles, R. Kimmel, et G. Sapiro. Geodesic active contours. *Int. J. Comput. Vis.*, 22(1) :61–79, 1997.

- A. Chakraborty, H. Staib, et J. Duncan. Deformable boundary finding in medical images by integrating gradient and region information. *IEEE Trans. Med. Imag.*, 15(6) :859–870, Décembre 1996.
- A. Chambolle et T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imag. Vis.*, 40(1) :120–145, 2011.
- T. Chan, S. Esedoglu, et M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics*, 66(5) :1632–1648, 2006.
- T. Chan et L. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10(2) :266–277, 2001.
- T. F. Chan, M. Moelich, et B. Sandberg. Some recent developments in variational image segmentation. Dans *Image processing based on partial differential equations*, pages 175–210. Springer, 2007.
- C. Chen, J. Leng, et G. Xu. A general framework of piecewise-polynomial Mumford-Shah model for image segmentation. Rapport Technique 13-50, UCLA C.A.M. Report, Septembre 2013a.
- D. Chen, M. Yang, et L. D. Cohen. Global minimum for a variant Mumford-Shah model with application to medical image segmentation. *Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization*, 1(1) :48–60, 2013b.
- Y. Chen et W. Guo. A local nonparametric model for simultaneous image segmentation and adaptive smooth. Rapport Technique 07-34, UCLA C.A.M. Report, Septembre 2007.
- J. Cheng et S. Foo. Dynamic directional gradient vector flow for snakes. *IEEE Trans. Image Process.*, 15(6) :1563–1571, Juin 2006.
- C. Chesnaud, P. Réfrégier, et V. Boulet. Statistical region snake-based segmentation adapted to different physical noise models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(11) :1145–1157, 1999.
- J. Coatrieux. A look at... shape and function from motion in medical imaging : Part I. *IEEE Eng. Med. Biol. Mag.*, 24(3) :29–31, 2005.
- J. Coatrieux. Shape and function from motion in medical imaging : part 2. *IEEE Eng. Med. Biol. Mag.*, 25(1) :6–21, 2006.
- J. P. Cocquerez et S. Philipp-Foliguet. *Analyse d'images : filtrage et segmentation*. Enseignement de la physique. Masson, Octobre 1995.
- B. Cohen et I. Dinstein. Motion estimation in noisy ultrasound images by maximum likelihood. Dans *Int.l Conf. Pattern Recognit.*, volume 3, pages 182–185 vol.3, 2000.
- B. Cohen et I. Dinstein. New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognit.*, 35(2) :455–463, Février 2002.
- L. Cohen et I. Cohen. Finite-element methods for active contour models and balloons for 2D and 3D images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(11) :1131–1147, 1993.
- L. Cohen et R. Kimmel. Global minimum for active contour models : A minimal path approach. *Int. J. Comput. Vis.*, 24(1) :57–78, 1997.
- L. D. Cohen. On active contour models and balloons. *CVGIP : Image Understanding*, 53(2) : 211–218, 1991.
- L. D. Cohen. Avoiding local minima for deformable curves in image analysis. *Curves and Surfaces with Applications in CAGD*, pages 77–84, 1997.

- L. D. Cohen, E. Bardinet, et N. Ayache. Reconstruction of digital terrain model with a lake. Dans *Proc. of SPIE'93, Geometric Methods in Computer Vision*, San Diego, California, USA, États-Unis, 1993. Available as INRIA Research Report 1824, December 1992.
- O. Colliot, O. Camara, et I. Bloch. Integration of fuzzy spatial relations in deformable models—Application to brain MRI segmentation. *Pattern Recognit.*, 39(8) :1401–1414, Août 2006.
- D. Comaniciu, X. Zhou, et S. Krishnan. Robust real-time myocardial border tracking for echocardiography : An information fusion approach. *IEEE Trans. Med. Imag.*, 23(7) :849–860, 2004.
- D. Cremers. Dynamical statistical shape priors for level set-based tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(8) :1262–1273, 2006.
- D. Cremers, M. Rousson, et R. Deriche. A review of statistical approaches to level set segmentation : Integrating color, texture, motion and shape. *Int. J. Comput. Vis.*, 72(2) :195–215, 2007.
- A. Criminisi, T. Sharp, et A. Blake. GeoS : Geodesic image segmentation. Dans D. A. Forsyth, P. H. S. Torr, et A. Zisserman, éditeurs, *ECCV (1)*, volume 5302 de *Lecture Notes in Computer Science*, pages 99–112. Springer, 2008. ISBN 978-3-540-88681-5.
- C. Darolt, A. Mertins, C. Bodensteiner, et U. G. Hofmann. Local region descriptors for active contours evolution. *IEEE Trans. Image Process.*, 17(12) :2275–2288, 2008.
- G. Demarcq, L. Mascarilla, M. Berthier, et P. Courtellemont. The color monogenic signal : Application to color edge detection and color optical flow. *J. Math. Imag. Vis.*, 40(3) :269–284, Juillet 2011.
- A. P. Dempster, N. M. Laird, et D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 34 :1–38, 1977.
- R. Deriche. Using Canny's criteria to derive a recursively implemented optimal edge detector. *Int. J. Comput. Vis.*, 1(2) :167–187, 1987.
- R. Donner, B. Micusik, G. Langs, L. Szumilas, P. Peloschek, K. Friedrich, et H. Bischof. Object localization based on markov random fields and symmetry interest points. Dans *MICCAI'07*, pages 460–468, Berlin, Heidelberg, 2007. Springer-Verlag.
- R. Duits, M. Felsberg, L. Florack, et B. Platel. α scale spaces on a bounded domain. Dans *Proc. of the 4th Int. conference on Scale space methods in computer vision*, Scale Space'03, pages 494–510, Berlin, Heidelberg, 2003. Springer-Verlag.
- R. Duits, L. Florack, J. De Graaf, et B. ter Haar Romeny. On the axioms of scale space theory. *J. Math. Imag. Vis.*, 20(3) :267–298, 2004.
- V. Duval, J.-F. Aujol, et Y. Gousseau. A bias-variance approach for the nonlocal means. *SIAM J. Imag. Sciences*, 4(2) :760–788, 2011.
- I. Dydenko, F. Jamal, O. Bernard, J. D'hooge, I. E. Magnin, et D. Friboulet. A level set framework with a shape and motion prior for segmentation and region tracking in echocardiography. *Med. Image Anal.*, 10(2) :162–177, Avril 2006.
- N. El-Zehiry et A. Elmaghraby. A graph cut based active contour without edges with relaxed homogeneity constraint. Dans *Int'l Conf. Pattern Recognit.*, pages 1–4, 2008.
- J. H. Elder et S. W. Zucker. Local scale control for edge detection and blur estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(7) :699–716, Juillet 1998.
- M. Erdt, S. Steger, et G. Sakas. Regmentation : A new view of image segmentation and registration. *Journal of Radiation Oncology Informatics*, 4(1) :1–23, 2012.

- A. Faisal et C. Pluempitiwiriawej. Active contour using local region-scalable force with expandable kernel. Dans *Int. Conf. on Information Science and Technology (ICIST)*, pages 18–24, 2012.
- S. Fazekas, T. Amiaz, D. Chetverikov, et N. Kiryati. Dynamic texture detection based on motion analysis. *Int. J. Comput. Vis.*, 82(1) :48–63, 2009.
- M. Feigin, N. Sochen, et B. Vemuri. Anisotropic α -kernels and associated flows. *SIAM J. on Imag. Sciences*, 3(4) :904–925, 2010.
- M. Felsberg, R. Duits, et L. Florack. The monogenic scale space on a rectangular domain and its features. *Int. J. Comput. Vis.*, 64 :187–201, 2005.
- M. Felsberg et G. Sommer. A new extension of linear signal processing for estimating local properties and detecting features. Dans *In 22. DAGM Symposium Mustererkennung*, pages 195–202, 2000.
- M. Felsberg et G. Sommer. The monogenic signal. *IEEE Trans. Signal Process.*, 49(49) :3136–3144, 2001a.
- M. Felsberg et G. Sommer. Scale adaptive filtering derived from the laplace equation. Dans *Proc. of the 23rd DAGM-Symposium on Pattern Recognition*, pages 124–131, London, UK, 2001b. Springer-Verlag. ISBN 3-540-42596-9.
- M. Felsberg et G. Sommer. The monogenic scale-space : A unifying approach to phase-based image processing in scale-space. *J. Math. Imag. Vis.*, 21(1) :5–26, 2004.
- J. Feng, Z. Cao, et Y. Pi. Multiphase SAR image segmentation with G^0 -statistical-model-based active contours. *IEEE Trans. Geosci. Remote Sens.*, 51(7) :4190–4199, 2013.
- S. Fidler, R. Mottaghi, A. Yuille, et R. Urtasun. Bottom-up segmentation for top-down detection. Dans *Proc. of IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2013.
- D. J. Fleet et A. D. Jepson. Stability of phase information. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(12) :1253–1268, 1993.
- D. Freedman et T. Zhang. Active contours for tracking distributions. *IEEE Trans. Image Process.*, 13(4) :518–526, 2004.
- W. T. Freeman et E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(9) :891–906, 1991.
- J. Freixenet, X. Muñoz, D. Raba, J. Martí, et X. Cufí. Yet another survey on image segmentation : Region and boundary information integration. Dans *European Conf. Comput. Vis.*, pages 408–422, 2002.
- D. Gabor. Theory of communication. *J. Inst. Elec. Eng.*, 93(26) :429–457, 1946.
- S. Geman et D. Geman. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(6) :721–741, 1984.
- J. Geusebroek, A. Smeulders, et J. van de Weijer. Fast anisotropic gauss filtering. *IEEE Trans. Image Process.*, 12(8) :938–943, 2003.
- R. C. Gonzalez, R. E. Woods, et S. L. Eddins. *Digital Image Processing Using MATLAB*. Gatesmark Publishing, 2nd édition, 2009. ISBN 9780982085400.
- N. Gordillo, E. Montseny, et P. Sobrevilla. State of the art survey on MRI brain tumor segmentation. *Magnetic Resonance Imaging*, 31(8) :1426 – 1438, 2013.

- S. Gorthi, V. Duay, X. Bresson, M. B. Cuadra, F. J. S. Castro, C. Pollo, A. S. Allal, et J.-P. Thiran. Active deformation fields : Dense deformation field estimation for atlas-based segmentation using the active contour framework. *Med. Image Anal.*, 15(6) :787 – 800, 2011.
- G. H. Granlund et H. Knutsson. *Signal processing for computer vision*. Kluwer Academic Publisher, 1995.
- Y. Gu, L.-L. Wang, et X.-C. Tai. A direct approach toward global minimization for multiphase labeling and segmentation problems. *IEEE Trans. Image Process.*, 21(5) :2399–2411, 2012.
- L. Guillot et C. Le Guyader. Extrapolation of vector fields using the infinity laplacian and with applications to image segmentation. Dans X.-C. Tai et al., éditeur, *SSVM*, volume 5567 de *Lect. Note Comput. Sci.*, pages 87–99. Springer Berlin Heidelberg, 2009.
- W. Guo et Y. Chen. Using non-parametric kernel to segment and smooth images simultaneously. Dans *IEEE Int. Conf. Image Proc.*, pages 217–220, 2006.
- Y. Guo et C.-C. Lu. Multi-modality image registration using mutual information based on gradient vector flow. Dans *Int.l Conf. Pattern Recognit.*, volume 3, pages 697–700, Los Alamitos, CA, USA, 2006.
- J. Hadamard. *Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques : leçons professées à l'Université Yale*. Mrs. Hepsa Ely Silliman memorial lectures. Hermann et cie, 1932.
- X. Han, C. Xu, et J. Prince. Fast numerical scheme for gradient vector flow computation using a multigrid method. *IET Image Proc.*, 1(1) :48–55, Mars 2007.
- R. M. Haralick. Digital step edges from zero crossing of second directional derivatives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(1) :58–68, 1984.
- M. Hassouna et A. Farag. On the extraction of curve skeletons using gradient vector flow. Dans *Int. Conf. Computer Vision, Brazil, Oct. 14-20*, pages 1–8, 2007.
- M. Heiler et C. Schnörr. Natural image statistics for natural image segmentation. *Int. J. Comput. Vis.*, 63(1) :5–19, 2005.
- T. Heimann et H.-P. Meinzer. Statistical shape models for 3D medical image segmentation : A review. *Med. Image Anal.*, 13(4) :543 – 563, 2009.
- F. Heitger. Feature detection using suppression and enhancement. Rapport Technique TR-163, Image Science Lab, 1995.
- F. Heitger, L. Rosenthaler, R. Von Der Heydt, E. Peterhans, et O. Kübler. Simulation of neural contour mechanisms : from simple to end-stopped cells. *Vision research*, 32(5) :963–981, 1992.
- A. Herbulot, S. Jehan-Besson, S. Duffner, M. Barlaud, et G. Aubert. Segmentation of vectorial image features using shape gradients and information measures. *J. Math. Imag. Vis.*, 25(3) : 365–386, 2006.
- L. S. Hibbard. Region segmentation using information divergence measures. *Med. Image Anal.*, 8(3) :233–244, 2004.
- M. Holtzman-Gazit, R. Kimmel, N. Peled, et D. Goldsher. Segmentation of thin structures in volumetric medical images. *IEEE Trans. Image Process.*, 15(2) :354–363, 2006.
- N. Houhou, J.-P. Thiran, et X. Bresson. Fast texture segmentation based on semi-local region descriptor and active contour. *Numerical Mathematics : Theory, Methods and Applications*, 2(4) : 445–468, 2009.

- J. E. Iglesias, M. R. Sabuncu, et K. V. Leemput. A unified framework for cross-modality multi-atlas segmentation of brain MRI. *Med. Image Anal.*, 17(8) :1181 – 1191, 2013.
- A. C. Jalba, M. H. F. Wilkinson, et J. B. T. M. Roerdink. CPM : A deformable model for shape recovery and segmentation based on charged particles. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(10) :1–16, 2004.
- S. Jehan-Besson, M. Barlaud, et G. Aubert. DREAM2S : deformable regions driven by an Eulerian accurate minimization method. *Int. J. Comput. Vis.*, 53(1) :45–70, 2003.
- J. A. Jensen. Field : A program for simulating ultrasound systems. *Med. Biol. Eng. Comput.*, 34 (Supplement 1, Part 1) :351–353, 1996.
- N. Joshi et M. Brady. Non-parametric mixture model based evolution of level sets and application to medical images. *Int. J. Comput. Vis.*, 88(1) :52–68, 2010.
- M. Jung, G. Peyré, et L. D. Cohen. Texture segmentation via non-local non-parametric active contours. Dans *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 74–88. Springer, 2011.
- M. Jung, G. Peyré, et L. D. Cohen. Nonlocal active contours. *SIAM J. Imag. Sciences*, 5(3) : 1022–1054, 2012.
- T. Kadir et M. Brady. Unsupervised non-parametric region segmentation using level sets. Dans *IEEE Int. Conf. Comput. Vis.*, pages 1267–1274 vol.2, 2003.
- M. Kass, A. Witkin, et D. Terzopoulos. Snakes : Active contour models. *Int. J. Comput. Vis.*, 1(4) : 321–331, 1988.
- V. Katkovnik. A new method for varying adaptive bandwidth selection. *IEEE Trans. Signal Process.*, 47(9) :2567–2571, 1999.
- V. Katkovnik, K. Egiazarian, et J. Astola. Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule. *J. Math. Imag. Vis.*, 16(3) :223–235, 2002.
- V. Katkovnik, K. Egiazarian, et J. Astola. *Local approximation techniques in signal and image processing*. Int. Society for Optical Eng., Washington, 2006. ISBN 0819460923.
- V. Katkovnik, A. Foi, K. Egiazarian, et J. Astola. Directional varying scale approximations for anisotropic signal processing. Dans *European Signal Proc. Conf.*, pages 101–104, Vienna, Austria, Septembre 2004.
- A. Katouzian, E. Angelini, S. Carlier, J. Suri, N. Navab, et A. Laine. A state-of-the-art review on segmentation algorithms in intravascular ultrasound (ivus) images. *IEEE Trans. Inf. Technol. Biomed.*, 16(5) :823–834, 2012.
- C. Kervrann et J. Boulanger. Optimal spatial adaptation for patch-based image denoising. *IEEE Trans. Image Process.*, 15(10) :2866–2878, 2006.
- J. Kim, J. Fisher, A. Yezzi, M. Çetin, et A. Willsky. A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE Trans. Image Process.*, 14 : 1486–1502, 2005.
- R. Kimmel. Geometric level set methods in imaging, vision, and graphics. Chapitre Fast edge integration, pages 59–77. New York : Springer, 2003.
- R. Kimmel et A. Bruckstein. Regularized laplacian zero crossings as optimal edge integrators. *Int. J. Comput. Vis.*, 53(3) :225–243, 2003.
- D. Kirsanov et S. J. Gortler. A discrete global minimization algorithm for continuous variational problems. *Harvard Computer Science, Tech. Rep. TR-14-04*, 2004.

- R. Klette et P. Zamperoni. *Handbook of image processing operators*. John Wiley & Sons, Inc., New York, NY, USA, 1996.
- H. Knutsson et M. Andersson. Loglets : Generalized quadrature and phase for local spatiotemporal structure estimation. Dans *In Proc. Scandinavian Conf. Image Anal.*, 2003a.
- H. Knutsson et M. Andersson. What's so good about quadrature filters. Dans *IEEE Int. Conf. Image Proc.*, pages 61–64, 2003b.
- G. Koepfler, C. Lopez, et J.-M. Morel. A multiscale algorithm for image segmentation by variational method. *SIAM journal on numerical analysis*, 31(1) :282–299, 1994.
- V. Kolmogorov et Y. Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. Dans *IEEE Int. Conf. Comput. Vis.*, volume 1, pages 564–571, 2005.
- V. Kolmogorov et R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2) :147–159, Février 2004.
- D. Kondermann, S. Abraham, G. Brostow, W. Förstner, S. Gehrig, A. Imiya, B. Jähne, F. Klose, M. Magnor, H. Mayer, R. Mester, T. Pajdla, R. Reulke, et H. Zimmer. On performance analysis of optical flow algorithms. Dans F. Dellaert, J.-M. Frahm, M. Pollefeys, L. Leal-Taixé, et B. Rosenhahn, éditeurs, *Outdoor and Large-Scale Real-World Scene Analysis*, volume 7474 de *Lecture Notes in Computer Science*, pages 329–355. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-34090-1.
- S. Konishi, A. L. Yuille, J. M. Coughlan, et S. C. Zhu. Statistical edge detection : learning and evaluating edge cues. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25 :57–74, 2003.
- P. Kovesi. *Invariant measures of image features from phase information*. PhD thesis, University of Western Australia, 1996.
- P. Kovesi. Symmetry and asymmetry from local phase. Dans *Tenth Australian joint conference on artificial intelligence*, volume 190. Citeseer, 1997.
- P. Kovesi. Image features from phase congruency. *Videre : J. Comput. Vis. Res.*, 1(3) :1–26, 1999.
- P. Kovesi. Phase congruency detects corners and edges. Dans *Australian Pattern Recognition Society Conf. : DICTA*, pages 309–318, 2003.
- W. M. Krueger et K. Phillips. The geometry of differential operators with application to image processing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(12) :1252–1264, 1989.
- P. Kube et P. Perona. Scale-space properties of quadratic feature detectors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(10) :987–999, 1996.
- S. Lankton et A. Tannenbaum. Localizing region-based active contours. *IEEE Trans. Image Process.*, 17(11) :2029–2039, 2008.
- Y.-N. Law, H. K. Lee, et A. Yip. A multiresolution stochastic level set method for Mumford-Shah image segmentation. *IEEE Trans. Image Process.*, 17(12) :2289–2300, 2008.
- T. M. Le et L. A. Vese. Additive & mutiplicative piecewise-smooth segmentation models in a functional minimization approach. *Contemporary Mathematics*, 445 :207–224, 2007.
- F. Lecellier, J. Fadili, S. Jehan-Besson, G. Aubert, M. Revenu, et E. Saloux. Region-based active contours with exponential family observations. *J. Math. Imag. Vis.*, 36(1) :28–45, 2010.
- M. J. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Sühling, P. Hunziker, et M. Unser. Spatio-temporal non-rigid registration for ultrasound cardiac motion estimation. *IEEE Trans. Med. Imag.*, 24(9), Septembre 2005.

- O. Lepskii. On a problem of adaptive estimation in gaussian white noise. *Theory of Probability & Its Applications*, 35(3) :454–466, 1990.
- B. Leroy, I. Herlin, et L. Cohen. Multi-resolution algorithm for active contour models. In *12th Int. Conf. Anal. Opt. System*, 15(11) :58–65, 1996.
- A. Levin et Y. Weiss. Learning to combine bottom-up and top-down segmentation. *Int. J. Comput. Vis.*, 81(1) :105–118, 2009.
- B. Li et S. T. Acton. Active contour external force using vector field convolution for image segmentation. *IEEE Trans. Image Process.*, 16(8) :2096–2106, 2007.
- C. Li, R. Huang, Z. Ding, J. C. Gatenby, et D. N. Metaxas. A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Trans. Image Process.*, 20(7) :2007–2016, 2011a.
- C. Li, C.-Y. Kao, J. C. Gore, et Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans. Image Process.*, 17(10) :1940–1949, 2008.
- C. Li, J. Liu, et M. D. Fox. Segmentation of external force field for automatic initialization and splitting of snakes. *Pattern Recognit.*, 38(11) :1947 – 1960, 2005.
- H. Li et A. Yezzi. Local or global minima : flexible dual-front active contours. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(1) :1–14, 2007.
- J. Li, W. Zuo, X. Zhao, et D. Zhang. An augmented lagrangian method for fast gradient vector flow computation. Dans *IEEE Int. Conf. Image Proc.*, pages 1525–1528, 2011b.
- J. Liang, T. McInerney, et D. Terzopoulos. United snakes. *Med. Image Anal.*, 10(2) :215–233, 2006.
- T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers : Netherlands, 1994.
- T. Lindeberg. Feature detection with automatic scale selection. *Int. J. Comput. Vis.*, 30(2) :79–116, 1998.
- T. Lindeberg. Scale-space. Dans e. Benjamin Wah, éditeur, *Encyclopedia of Computer Science and Engineering*, volume IV, pages 2495–2504. John Wiley and Sons, Hoboken, New Jersey, Septembre 2008.
- T. Lindeberg. Generalized gaussian scale-space axiomatics comprising linear scale-space, affine scale-space and spatio-temporal scale-space. *J. Math. Imag. Vis.*, 40(1) :36–81, 2011.
- M. Linguraru, N. Vasilyev, G. Marx, W. Twaretzky, P. Del Nido, et R. Howe. Fast block flow tracking of atrial septal defects in 4D echocardiography. *Med. Image Anal.*, 12(4) :397–412, 2008.
- G. Liu, G.-S. Xia, W. Yang, et M.-S. Liao. SAR image segmentation via non-local active contours. Dans *APSAR*, pages 1–4, Tsukuba, Japon, Septembre 2013.
- J. Liu et H. Zhang. Image segmentation using a local GMM in a variational framework. *J. Math. Imag. Vis.*, pages 1–16, 2013.
- M. Lorenzo-Valdes, G. Sanchez-Ortiz, A. Elkington, R. H. Mohiaddin, et D. Rueckert. Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. *Med. Image Anal.*, 8(3) :255–265, 2004.
- C. Lu, S. Chelikani, X. Papademetris, J. P. Knisely, M. F. Milosevic, Z. Chen, D. A. Jaffray, L. H. Staib, et J. S. Duncan. An integrated approach to segmentation and nonrigid registration for application in image-guided pelvic radiotherapy. *Medical Image Analysis*, 15(5) :772 – 785, 2011.

- S. Mahmoodi et B. Sharif. Contour evolution scheme for variational image segmentation and smoothing. *IET Image Process.*, 1(3) :287–294, 2007.
- R. Malladi, J. Sethian, et B. Vemuri. Shape modeling with front propagation : A level set approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(2) :158–175, 1995.
- R. Malladi, J. Sethian, et B. Vemuri. A fast level set based algorithm for topology-independent shape modeling. *J. Math. Imaging Vis.*, 6(2-3) :269–289, 1996.
- S. G. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- A.-R. Mansouri. Region tracking via level set PDEs without motion computation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7) :947–961, 2002.
- D. Marr et E. Hildreth. Theory of edge detection. Dans *Proc of the Royal Society London B*, 207, pages 187–21, 1980.
- J. L. Marroquin, J. E. Figueroa, et M. Servin. Robust quadrature filters. *J. Opt. Soc. Am. A*, 14(4) : 779–791, Apr 1997.
- D. Martin, C. C. Fowlkes, et J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(5) :530–549, 2004a.
- P. Martin, P. Réfrégier, F. Galland, et F. Guérault. Nonparametric statistical snake based on the minimum stochastic complexity. *IEEE Trans. Image Process.*, 15(9) :2762–2770, 2006.
- P. Martin, P. Réfrégier, F. Goudail, et F. Guérault. Influence of the noise model on level set active contour segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(6) :799–803, Juin 2004b.
- W. Mcilhagga. The canny edge detector revisited. *Int. J. Comput. Vis.*, 91(3) :251–261, Février 2011.
- T. McInerney et D. Terzopoulos. Deformable models in medical image analysis : A survey. *Med. Image Anal.*, 1(2) :91–108, 1996.
- M. Mellor, B.-W. Hong, et M. Brady. Locally rotation, contrast, and scale invariant descriptors for texture analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(1) :52–61, 2008.
- O. Michailovich, Y. Rathi, et A. Tannenbaum. Image segmentation using active contours driven by the Bhattacharyya gradient flow. *IEEE Trans. Image Process.*, 16(11) :2787–2801, 2007.
- M. Mignotte, J. Meunier, et J.-C. Tardif. Endocardial boundary estimation and tracking in echocardiographic images using deformable template and markov random fields. *Pattern Analysis & Applications*, 4(4) :256–271, 2001.
- I. Mikić, S. Krucinski, et J. D. Thomas. Segmentation and tracking in echocardiographic sequences : Active contours guided by optical flow estimates. *IEEE Trans. Med. Imag.*, 17(2) : 274–284, Avril 1998.
- J. Mille. Narrow band region-based active contours and surfaces for 2D and 3D segmentation. *Comput. Vision Image Understanding*, 113(9) :946–965, 2009.
- A. Mitiche et I. Ben Ayed. *Variational and Level Set Methods in Image Segmentation*. Springer topics in signal processing. Springer, 2010. ISBN 9783642153525.
- A. Mitiche et H. Sekkati. Optical flow 3D segmentation and interpretation : A variational method with active curve evolution and level sets. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11) :1818–1829, 2006.
- J. M. Morel et S. Solimini. Segmentation of images by variational methods : A constructive approach. *Rev. Mat. Univ. Complut. Madrid*, 1(1-3) :169–182, 1988.

- J.-M. Morel et S. Solimini. *Variational Methods for Image Segmentation : with seven image processing experiments*. Progress in nonlinear differential equations and their applications. Boston, MA : Birkhauser, 1995. ISBN 978-1-4684-0569-9.
- M. Morrone, A. Navangione, et D. Burr. An adaptive approach to scale selection for line and edge detection. *Pattern Recognit. Lett.*, 16(7) :667 – 677, 1995.
- M. C. Morrone et R. A. Owens. Feature detection from local energy. *Pattern Recognit. Lett.*, 6(5) : 303–313, 1987.
- M. Mulet-Parada et J. Noble. 2D+T acoustic boundary detection in echocardiography. *Med. Image Anal.*, 4(1) :21–30, 2000.
- D. Mumford. Bayesian rationale for the variational formulation. Dans B. Haar Romeny, éditeur, *Geometry-Driven Diffusion in Computer Vision*, volume 1 de *Computational Imaging and Vision*, pages 135–146. Springer Netherlands, 1994. ISBN 978-90-481-4461-7.
- D. Mumford et J. Shah. Boundary detection by minimizing functionals. Dans *Proc. of IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 1985.
- D. Mumford et J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun Pure Appl Math*, 42(5) :577–685, 1989.
- S. Nadarajah. Statistical distributions of potential interest in ultrasound speckle analysis. *Phys. Med. Biol.*, 52 :213–227, 2007.
- J. Ng et A. A. Bharath. Steering in scale space to optimally detect image structures. Dans *Computer Vision-ECCV*, pages 482–494. Springer, 2004.
- K. Ni, X. Bresson, T. Chan, et S. Esedoglu. Local histogram based segmentation using the waserstein distance. *Int. J. Comput. Vis.*, 84(1) :97–111, 2009.
- C. Nieuwenhuis, E. Töppe, et D. Cremers. A survey and comparison of discrete and continuous multi-label optimization approaches for the potts model. *J. Math. Imag. Vis.*, 104(3) :223–240, 2013.
- J. Ning, C. Wu, S. Liu, et S. Yang. NGVF : An improved external force field for active contour model. *Pattern Recognit. Lett.*, 28(1) :58–63, Janvier 2007.
- A. Noble et D. Boukerroui. Ultrasound image segmentation : A survey. *IEEE Trans. Med. Imag.*, 25(8) :987–1010, 2006.
- S. Olabarriaga et A. Smeulders. Interaction in the segmentation of medical images : A survey. *Med. Image Anal.*, 5(2) :127 – 142, 2001.
- C. F. Olson. Adaptive-scale filtering and feature detection using range data. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(9) :983–991, 2000.
- A. Oppenheim et J. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5) : 529–541, Mai 1981.
- S. Osher et N. Paragolis. *Geometric Level Set Methods in Imaging Vision and Graphics*. Springer Verlag, 2003.
- S. Osher et J. A. Sethian. Fronts propagating with curvature-dependent speed : Algorithms based on Hamilton-Jacobi formulations. *J Comput Phys*, 79(1) :12–49, 1988.
- R. Owens, S. Venkatesh, et J. Ross. Edge detection is a projection. *Pattern Recognit. Lett.*, 9(4) : 233–244, 1989.
- G. Papari et N. Petkov. Edge and line oriented contour detection : State of the art. *Image Vis. Comput.*, 29(2-3) :79–103, 2011.

- T. N. Pappas. An adaptive clustering algorithm for image segmentation. *IEEE Trans. Signal Process.*, 40(4) :901–914, 1992.
- N. Paragios. A variational approach for the segmentation of the left ventricle in cardiac image analysis. *Int. J. Comput. Vis.*, 50 :345–362, 2002.
- N. Paragios. A level set approach for shape-driven segmentation and tracking of the left ventricle. *IEEE Trans. Med. Imag.*, 22(6) :773–776, 2003.
- N. Paragios, Y. Chen, et O. Faugeras. *Handbook of Mathematical Models in Computer Vision*. Springer, (1st edition), Octobre 2005.
- N. Paragios et R. Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(3) :266–280, 2000.
- N. Paragios et R. Deriche. Geodesic active regions : A new framework to deal with frame partition problems in computer vision. *J. Visual Commun. Image Representation*, 13(1) :249–268, 2002.
- N. Paragios et R. Deriche. Geodesic active regions and level set methods for motion estimation and tracking. *Computer Vision and Image Understanding*, 97(3) :259–282, 2005.
- N. Paragios, O. Mellina-Gottardo, et V. Ramesh. Gradient vector flow fast geometric active contours. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(3) :402–407, Mars 2004.
- H. Park, P. Bland, et C. Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans. Med. Imag.*, 22(4) :483–492, Février 2003.
- E. Pauwels, L. Van Gool, P. Fiddelaers, et T. Moons. An extended class of scale-invariant and recursive scale space filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(7) :691–701, 1995.
- F. Pellegrino, W. Vanzella, et V. Torre. Edge detection revisited. *IEEE Trans. Syst., Man, Cybern. B*, 34 :1500–1518, 2004.
- B. Peng, L. Zhang, et D. Zhang. A survey of graph theoretical approaches to image segmentation. *Pattern Recognit.*, 46(3) :1020 – 1038, 2013.
- P. Perona et J. Malik. Detecting and localizing edges composed of steps, peaks and roofs. Dans *IEEE Int. Conf. Comput. Vis.*, pages 52–57, 1990a.
- P. Perona et J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12 :629–639, Juillet 1990b.
- G. Peyre, J. Fadili, et J. Rabin. Wasserstein active contours. Dans *IEEE Int. Conf. Image Proc.*, pages 2541–2544, 2012.
- B. Picinbono. On instantaneous amplitude and phase of signals. *IEEE Trans. Signal Process.*, 45 (3) :552–560, Mars 1997.
- J. Piovan et T. Papadopoulos. Local statistic based region segmentation with automatic scale selection. Dans *European Conf. Comput. Vis.*, pages 486–499, Marseille, France, 2008.
- J. Piovan, M. Rousson, et T. Papadopoulos. Efficient segmentation of piecewise smooth images. Dans F. Sgallari, A. Murli, et N. Paragios, éditeurs, *Scale Space and Variational Methods in Computer Vision*, volume 4485 de *Lecture Notes in Computer Science*, pages 709–720. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-72822-1.
- T. Pock, D. Cremers, H. Bischof, et A. Chambolle. An algorithm for minimizing the Mumford-Shah functional. Dans *IEEE Int. Conf. Comput. Vis.*, pages 1133–1140, 2009.
- T. Poggio, H. Voorhees, et A. Yuille. A regularized solution to edge detection. *J. Complexity*, 4 (2) :106–123, 1988.

- L. Qin, C. Zhu, Y. Zhao, H. Bai, et H. Tian. Generalized gradient vector flow for snakes : New observations, analysis, and improvement. *IEEE Trans. Circuits and Systems for Video Technology*, 23(5) :883–897, 2013.
- K. Rajpoot, V. Grau, et J. A. Noble. Local-phase based 3d boundary detection using monogenic signal and its application to real-time 3-D echocardiography images. Dans *IEEE Int. Symp. Biomed. Imag.*, pages 783–786, 2009.
- N. Ray et S. T. Acton. Motion gradient vector flow : An external force for tracking rolling leukocytes with shape and size constrained active contours. *IEEE Trans. Med. Imag.*, 23(12) : 1466–1478, 2004.
- D. Reisfeld. The constrained phase congruency feature detector : Simultaneous localisation, classification and scale determination. *Pattern Recognit. Lett.*, 17 :1161–69, 1996.
- B. Robbins et R. A. Owens. 2D feature detection via local energy. *Image Vis. Comput.*, 15(5) : 353–368, 1997.
- T. Rohlfing, R. Brandt, R. Menzel, et C. R. M. Jr. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage*, 21(4) :1428 – 1442, 2004.
- R. Ronfard. Region-based strategies for active contour models. *Int. J. Comput. Vis.*, 13(2) :229–251, 1994.
- C. Ronse. On idempotence and related requirements in edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(5) :484–491, Mai 1993.
- C. Ronse. The phase congruence model for edge detection in multi-dimensional pictures. Report 97/16, LSIT, Université Louis Pasteur, Strasbourg, France, Octobre 1997.
- M. Rousson et R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. Rapport Technique 4515, INRIA, Juillet 2002.
- D. Rueckert, M. Lorenzo-Valdes, R. Chandrashekara, G. Sanchez-Ortiz, et R. Mohiaddin. Non-rigid registration of cardiac MR : Application to motion modelling and atlas-based segmentation. Dans *IEEE Int. Symposium on Biomedical Imaging*, pages 481–484, Washington DC, USA, 2002.
- C. Samson, L. Blanc-Féraud, G. Aubert, et J. Zerubia. A level set model for image classification. *Int. J. Comput. Vis.*, 40(3) :187–197, 2000.
- S. Sarkar et K. L. Boyer. On optimal infinite impulse response edge detection filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(11) :1154–1171, 1991.
- A. Sarti, C. Corsi, E. Mazzini, et C. Lamberti. Maximum likelihood segmentation of ultrasound images with Rayleigh distribution. *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, 52(6) :947–960, 2005.
- A. Sawatzky, D. Tenbrinck, X. Jiang, et M. Burger. A variational framework for region-based segmentation incorporating physical noise models. *J. Math. Imag. Vis.*, 47(3) :179–209, 2013.
- A. Sedlazeck. *Local Feature Detection by Higher Order Riesz Transforms on Images*. PhD thesis, Christian-Albrechts-Universität zu Kiel, 2008.
- J. Shen et S. Castan. An optimal linear operator for step edge detection. *CVGIP : Graphical Models and Image Processing*, 54(2) :112–133, 1992.
- J. Shotton, J. Winn, C. Rother, et A. Criminisi. Textonboost for image understanding : Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *Int. J. Comput. Vis.*, 81(1) :2–23, 2009.

- E. Simoncelli et H. Farid. Steerable wedge filters for local orientation analysis. *IEEE Trans. Image Process.*, 5(9) :1377–1382, 1996.
- E. Simoncelli, W. Freeman, E. Adelson, et D. Heeger. Shiftable multiscale transforms. *IEEE Trans. Inf. Theory*, 38(2) :587–607, 1992.
- A. Singh et P. Allen. Image-flow computation : An estimation-theoretic framework and a unified perspective. *Computer Vision, Graphics and Image Processing : Image Understanding*, 56(2) :152–177, Septembre 1992.
- G. Sommer et D. Zang. Parity symmetry in multi-dimensional signals. *Comm. Pure Appl. Anal.*, 6(3) :829–852, 2007.
- A. Stein et M. Hebert. Using spatio-temporal patches for simultaneous estimation of edge strength, orientation, and motion. Dans *Computer Vision and Pattern Recognition Workshop*, 2006. CVPRW '06. Conf. on, pages 19–19, 2006. To read.
- J. Stöttinger, R. Donner, L. Szumilas, et A. Hanbury. Evaluation of gradient vector flow for interest point detection. Dans G. Bebis et al., éditeur, *Advances in Visual Computing*, volume 5358 de *Lect. Note Comput. Sci*, pages 338–348. Springer Berlin Heidelberg, 2008.
- B. Sumengen et B. S. Manjunath. Multi-scale edge detection and image segmentation. Dans *European Signal Proc. Conf.*, Septembre 2005.
- D. Sun, S. Roth, et M. J. Black. A quantitative analysis of current practices in optical flow estimation and the principles behind them. Rapport Technique CS-10-03, Brown University, Department of Computer Science, Janvier 2013.
- R. Szeliski. *Computer Vision : Algorithms and Applications*. Springer-Verlag New York, 2010.
- W. Tao. Iterative narrowband-based graph cuts optimization for geodesic active contours with region forces (GACWRF). *IEEE Trans. Image Process.*, 21(1) :284–296, 2012.
- W. Tao, F. Chang, L. Liu, H. Jin, et T. Wang. Interactively multiphase image segmentation based on variational formulation and graph cuts. *Pattern Recognit.*, 43(10) :3208–3218, 2010.
- W. Tao et X.-C. Tai. Multiple piecewise constant with geodesic active contours (mpc-gac) framework for interactive image segmentation using graph cut optimization. *Image Vis. Comput.*, 29(8) :499–508, 2011.
- A. Tarachandani et D. Boltz. Review of the basic image processing and segmentation techniques for biological images. *J Imaging Sci Technol*, 50(3) :233 – 242, 2006.
- C. Tauber, H. Batatia, et A. Ayache. Une méthode d’initialisation automatique pour contour actif paramétrique, application aux images échographiques. Dans *GRETSI*, volume 1, pages 521–525, Louvain-la-Neuve, Belgique, 2005a.
- P. Tauber, H. Batatia, et A. Ayache. A robust active contour initialization and gradient vector flow for ultrasound image segmentation. Dans *IAPR, Tsukuba, Japan*. Prof. Katsushi Ikeuchi, Mai 2005b.
- V. Tavakoli et A. A. Amini. A survey of shaped-based registration and segmentation techniques for cardiac images. *Computer Vision and Image Understanding*, 117(9) :966 – 989, 2013.
- C. Tobon-Gomez, M. D. Craene, K. McLeod, L. Tautz, W. Shi, A. Hennemuth, A. Prakosa, H. Wang, G. Carr-White, S. Kapetanakis, A. Lutz, V. Rasche, T. Schaeffter, C. Butakoff, O. Frieman, T. Mansi, M. Sermesant, X. Zhuang, S. Ourselin, H.-O. Peitgen, X. Pennec, R. Razavi, D. Rueckert, A. Frangi, et K. Rhode. Benchmarking framework for myocardial tracking and deformation algorithms : An open access database. *Med. Image Anal.*, 17(6) :632 – 648, 2013.

- A. Tsai, A. Yezzi Jr, et A. S. Willsky. Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Trans. Image Process.*, 10(8) :1169–1186, 2001.
- M. Unser. Multigrid adaptive image processing. Dans *IEEE Int. Conf. Image Proc.*, volume 1, pages 49–52 vol.1, 1995.
- M. Unser, D. Sage, et D. Van De Ville. Multiresolution monogenic signal analysis using the Riesz-Laplace wavelet transform. *IEEE Trans. Image Process.*, 18(11) :2402–2418, Novembre 2009.
- M. Unser et D. Van De Ville. Wavelet steerability and the higher-order Riesz transform. *IEEE Trans. Image Process.*, 19(3) :636–652, Mars 2010.
- C. Vazquez, A.-R. Mansouri, et A. Mitiche. Approximation of images by basis functions for multiple region segmentation with level sets. Dans *IEEE Int. Conf. Image Proc.*, volume 1, pages 549–552, 2004.
- C. Vazquez, A. Mitiche, et R. Laganier. Joint multiregion segmentation and parametric estimation of image motion by basis function representation and level set evolution. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(5) :782–793, 2006.
- B. C. Vemuri, Y. C. J. Ye, et C. M. Leonard. Image registration via level-set motion : Applications to atlas-based segmentation. *Med. Image Anal.*, 7(1) :1–20, Mars 2003.
- S. Venkatesh et R. Owens. On the classification of image features. *Pattern Recognit. Lett.*, 11(5) : 339–349, 1990.
- L. Vese et T. Chan. A multiphase level set framework for image segmentation using the Mumford-Shah functional. *Int. J. Comput. Vis.*, 50(3) :271–293, 2002.
- E. Vidholm, P. Sundqvist, et I. Nyström. Accelerating the computation of 3D gradient vector flow fields. Dans *Int. Conf. Pattern Recognit.*, volume 3, pages 677–680, Los Alamitos, CA, USA, 2006.
- R. G. von Gioi, J. Jakubowicz, J.-M. Morel, et G. Randall. Lsd : A fast line segment detector with a false detection control. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32 :722–732, 2010.
- U. Vovk, F. Pernus, et B. Likar. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Trans. Med. Imag.*, 26(3) :405–421, 2007.
- C. Wachinger, T. Klein, et N. Navab. Locally adaptive nakagami-based ultrasound similarity measures. *Ultrasonics*, 52(4) :547–554, Apr 2012.
- H. Wang et A. Amini. Cardiac motion and deformation recovery from MRI : A review. *IEEE Trans. Med. Imag.*, 31(2) :487–503, 2012.
- L. Wang, L. He, A. Mishra, et C. Li. Active contours driven by local Gaussian distribution fitting energy. *Signal Process.*, 89 :2435–2447, 2009a.
- L. Wang, C. Li, Q. Sun, D. Xia, et C.-Y. Kao. Active contours driven by local and global intensity fitting energy with application to brain MR image segmentation. *Comput. Med. Imag. Graph.*, 33(7) :520–531, 2009b.
- Y. Wang, W. Chen, T. Yu, et Y. Zhang. Hessian based image structure adaptive gradient vector flow for parametric active contours. Dans *IEEE Int. Conf. Image Proc.*, pages 649–652, 2010.
- N. Werghi. Segmentation and modeling of full human body shape from 3-D scan data : A survey. *IEEE Trans. Syst., Man, Cybern.*, 37(6) :1122–1136, 2007.
- M. Werlberger, M. Unger, T. Pock, et H. Bischof. Efficient minimization of the non-local potts model. Dans *Scale Space and Variational Methods in Computer Vision*, pages 314–325. Springer, 2012.

- L. Wietzke, O. Fleischmann, et G. Sommer. 2D image analysis by generalized Hilbert transforms. Dans *Conformal Space, in ECCV, LNCS*. Springer-Verlag, 2008.
- L. Wietzke et G. Sommer. The 2D analytic signal. Rapport Technique o802, Kiel University, 2008.
- A. P. Witkin. Scale-space filtering. Dans *Proc. of the Eighth Int. joint conference on Artificial intelligence - Volume 2, IJCAI'83*, pages 1019–1022, San Francisco, CA, USA, 1983. Morgan Kaufmann Publishers Inc.
- B. Wu et Y. Yang. Local- and global-statistics-based active contour model for image segmentation. *Mathematical Problems in Eng.*, (ID 791958) :16, 2012.
- X. Xie. Active contouring based on gradient vector interaction and constrained level set diffusion. *IEEE Trans. Image Process.*, 19(1) :154–164, 2010.
- X. Xie et M. Mirmehdi. Mac : Magnetostatic active contour model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(4) :632–646, 2008.
- C. Xu et J. Prince. Generalized gradient vector flow external forces for active contours. *Signal Process.*, 71 :131–139, 1998a.
- C. Xu et J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.*, 7(3) : 359–369, 1998b.
- L. Xu, J. Chen, et J. Jia. A segmentation based variational model for accurate optical flow estimation. Dans *Computer Vision–ECCV 2008*, pages 671–684. Springer, 2008.
- N. Xu, N. Ahuja, et R. Bansal. Object segmentation using graph cuts based active contours. *Comput. Vision Image Understanding*, 107(3) :210–224, 2007.
- Q. Yang et D. Boukerroui. Optimal spatial adaptation for local region-based active contours : An intersection of confidence intervals approach. Dans *VISIGRAPP*, pages 87–93, Algarve, Portugal, Mars 2011.
- R. Yang, M. Mirmehdi, et X. Xie. A charged active contour based on electrostatics. Dans *Advanced Concepts for Intelligent Vision Systems*, pages 173–184. Springer-Verlag LNCS 4179, Septembre 2006.
- Y. Yang, S. Hallman, D. Ramanan, et C. C. Fowlkes. Layered object models for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(9) :1731–1743, 2012.
- Y. Yang, C. Li, C.-Y. Kao, et S. Osher. Split bregman method for minimization of region-scalable fitting energy for image segmentation. Rapport Technique 10-57, UCLA C.A.M. Report, Septembre 2010.
- M.-H. Yaou et W.-T. Chang. Fast surface interpolation using multiresolution wavelet transform. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(7) :673–688, 1994.
- A. Yezzi, A. Tsai, et A. Willsky. A fully global approach to image segmentation via coupled curve evolution equations. *J. Visual Commun. Image Representation*, 13(1-2) :195–216, 2002.
- R. Yildizoglu, J.-F. Aujol, et N. Papadakis. Active contours without level sets. Dans *IEEE Int. Conf. Image Proc.*, page 4, Orlando, Florida, USA, Septembre 2012.
- Y. Yue, J. Clark, et D. Khoury. Speckle tracking in intracardiac echocardiography for the assessment of myocardial deformation. *IEEE Trans. Biomed. Eng.*, 56(2) :416–425, 2009.
- X. Zeng, W. Chen, et Q. Peng. Efficiently solving the piecewise constant Mumford-Shah model using graph cuts. *Technical Report Zhejiang University*, 2006.

- Y. Zeng, D. Samaras, W. Chen, et Q. Peng. Topology cuts : A novel min-cut/max-flow algorithm for topology preserving segmentation in N-D images. *Comput. Vision Image Understanding*, 112(1) :81–90, 2008.
- D. Zhang et G. Lu. Segmentation of moving objects in image sequence : A review. Dans *Circuits, Systems and Signal Process*, pages 143–183, 2001.
- H. Zhang, Y. Chen, et J. Shi. Nonparametric image segmentation using Rényi’s statistical dependence measure. *J. Math. Imag. Vis.*, 44(3) :330–340, 2012.
- Y. Zhang, M. Brady, et S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.*, 20(1) :45–57, Janvier 2001.
- Q. Zheng, E. Dong, Z. Cao, W. Sun, et Z. Li. Modified localized graph cuts based active contour model for local segmentation with surrounding nearby clutter and intensity inhomogeneity. *Signal Process.*, 93(4) :961 – 966, 2013.
- S. K. Zhou, B. Georgescu, D. Comaniciu, et J. Shao. Boostmotion : Boosting a discriminative similarity function for motion estimation. Dans *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 1761–1768. IEEE, 2006.
- X. S. Zhou, D. Comaniciu, et A. Gupta. An information fusion framework for robust shape tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(1) :115–129, 2005.
- G. Zhu, S. Zhang, Q. Zeng, et C. Wang. Gradient vector flow active contours with prior directional information. *Pattern Recognit. Lett.*, 31(9) :845–856, 2010.
- S. C. Zhu et A. Yuille. Region competition : unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(9) :884–900, 1996.
- D. Ziou. Line detection using an optimal IIR filter. *Pattern Recognit.*, 24(6) :465–478, 1991.
- D. Ziou et S. Tabbone. Edge detection techniques : An overview. *Pattern Recognit. Image Anal.*, 8 : 537–559, 1998.

Troisième partie

Principales publications

On the choice of band-pass quadrature filters

Erratum

Djamal Boukerroui

Université de Technologie de Compiègne - CNRS, Heudiasyc UMR 7253
BP 20529 - 60205 Compiègne Cedex, France
djamal.boukerroui@hds.utc.fr

August 13, 2013

1 Phase definition

The phase as given by equation (5) in [1] and in [2] is wrong. Phase is the argument of the complex signal $f_A(x)$ and is 2π periodic. Thus, the four-quadrant inverse tangent has to be used:

$$\phi(x) = \arg(f_A(x)) = \text{atan2}(f_{\mathcal{H}}(x), f(x)) \quad .$$

Also, equation (7), giving the estimated local phase should be corrected accordingly, leading to,

$$\hat{\phi}(x) = \text{atan2}(f_o(x) \otimes f(x), f_e(x) \otimes f(x)) \quad .$$

2 Gaussian derivative filter

The given normalisation constant (eq. 34) and the peak tuning frequency (eq. 35) in [1, 2] are for the following definition of the Gaussian derivative quadrature filters in the frequency domain:

$$G_d(\omega) = \begin{cases} n_c \omega^a \exp[-\frac{1}{2}\sigma^2\omega^2], & \text{if } \omega \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

where $a \in \mathbb{R}$ and $a \geq 1$ (notice the extra $\frac{1}{2}$ coefficient in the exponential function).

For completeness, we recall the two equations bellow. The tuning frequency of the filter given above, is given by:

$$\omega_0 = \frac{\sqrt{a}}{\sigma} \quad , \quad (2)$$

and the unit energy normalisation constant n_c is given by:

$$\|g_d(x)\| = \frac{1}{2\pi} \|G_d(\omega)\| = 1 \quad \Rightarrow \quad n_c = 2 \frac{\sqrt{\pi} \sigma^{(a+\frac{1}{2})}}{\sqrt{\Gamma(a+\frac{1}{2})}} . \quad (3)$$

The correct tuning frequency and normalisation constant for the definition of the gaussian filter given in [1, 2], namely:

$$G_d(\omega) = \begin{cases} n_c \omega^a \exp[\sigma^2 \omega^2], & \text{if } \omega \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

can be obtained easily by replacing σ by $\sqrt{2}\sigma$ in eqs. (2) and (3). Leading to,

$$\omega_0 = \frac{1}{\sigma} \sqrt{\frac{a}{2}} , \quad (5)$$

and

$$\|g_d(x)\| = \frac{1}{2\pi} \|G_d(\omega)\| = 1 \quad \Rightarrow \quad n_c = 2^{\frac{2a+5}{4}} \frac{\sqrt{\pi} \sigma^{(a+\frac{1}{2})}}{\sqrt{\Gamma(a+\frac{1}{2})}} . \quad (6)$$

3 Phase signature

Phase signature of features as given in the legend of figure 1 “A local phase of $-\frac{\pi}{2}$, $+\frac{\pi}{2}$, 0 and π corresponds to a valley, a roof(line/bar), a negative step edge and a positive step edge respectively.” should be replaced by “A local phase of $\frac{\pi}{2}$, $-\frac{\pi}{2}$, 0 and π corresponds to a negative step edge, a positive step edge, a roof(line/bar) and a valley respectively.”

We mean by ‘a negative step edge’ an edge crossing from positive to negative values.

References

- [1] D. Boukerroui, J. A. Noble, and M. Brady, “On the choice of band-pass quadrature filters,” *J. Math. Imag. Vision*, vol. 21, no. 1, pp. 53–80, 2004.
- [2] —, *Frontiers in Robotics Research*, Max A. Denket ed. Nova Publishers, 2006, ch. On the Selection of Band-Pass Quadrature Filters, pp. 67–111.



On the Choice of Band-Pass Quadrature Filters

DJAMAL BOUKERROUI

Medical Vision Laboratory, Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK; HEUDIASYC, UMR CNRS 65999, Université de Technologie de Compiègne, BP 20529-60205 Compiègne Cedex, France

djamal.boukerroui@hds.utc.fr

J. ALISON NOBLE AND MICHAEL BRADY

Medical Vision Laboratory, Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK

noble@robots.ox.ac.uk

jmb@robots.ox.ac.uk

Abstract. Band-pass quadrature filters are extensively used in computer vision to estimate information from images such as: phase, energy, frequency and orientation,¹ possibly at different scales and utilise this in further processing-tasks. The estimation is intrinsically noisy and depends critically on the choice of the quadrature filters. In this paper, we first study the mathematical properties of the quadrature filter pairs most commonly seen in the literature and then consider some new pairs derived from the classical feature detection literature. In the case of feature detection, we present the first attempt to design a quadrature pair based on filters derived for optimal edge/line detection. A comparison of the filters is presented in terms of feature detection performance, wherever possible, in the sense of Canny and in terms of phase stability. We conclude with remarks on how our analysis can aid in the choice of a filter pair for a given image processing task.

Keywords: phase information, quadrature filter, feature detection, monogenic signal

1. Introduction

In computer vision and image analysis, arguably the most classical processing tools are amplitude-based techniques, both in the spatial and frequency domains. Consequently, throughout the history of digital image processing, smoothing and differentiation have been subjects of intense study. A variety of optimal differential operators have been proposed to solve different computer vision problems (see for example: [8, 13, 14, 37, 53, 63, 68, 69, 78, 79] for edge and line detection and [2, 16, 71] for optical flow estimation). Differentiation is highly sensitive to noise, but can be reduced/avoided using an appropriate scale selection for the smoothing function. However, amplitude-based measurements are sensitive to smooth shading and

lighting variations. In particular, this can be a source of serious error, for example for differential optical flow estimation. Moreover, an image contains many types of edges (not only step-edges). For instance, for edge and line detection, one might, for example, combine the outputs of a step detector and line detector. However, this is not a perfect solution since it leads to edge duplication (see for e.g. [66, 79]). Furthermore, especially for medical images such as X-ray mammograms [32], the interesting features in the image often do not have the idealised edge shape; here, we do not only mean blur degradation but also feature combination (for example feature number 4 in Fig. 1 is a combination of a line with a positive step edge). These may considerably reduce the optimality of the filter and lead to unpredictable multiple responses.

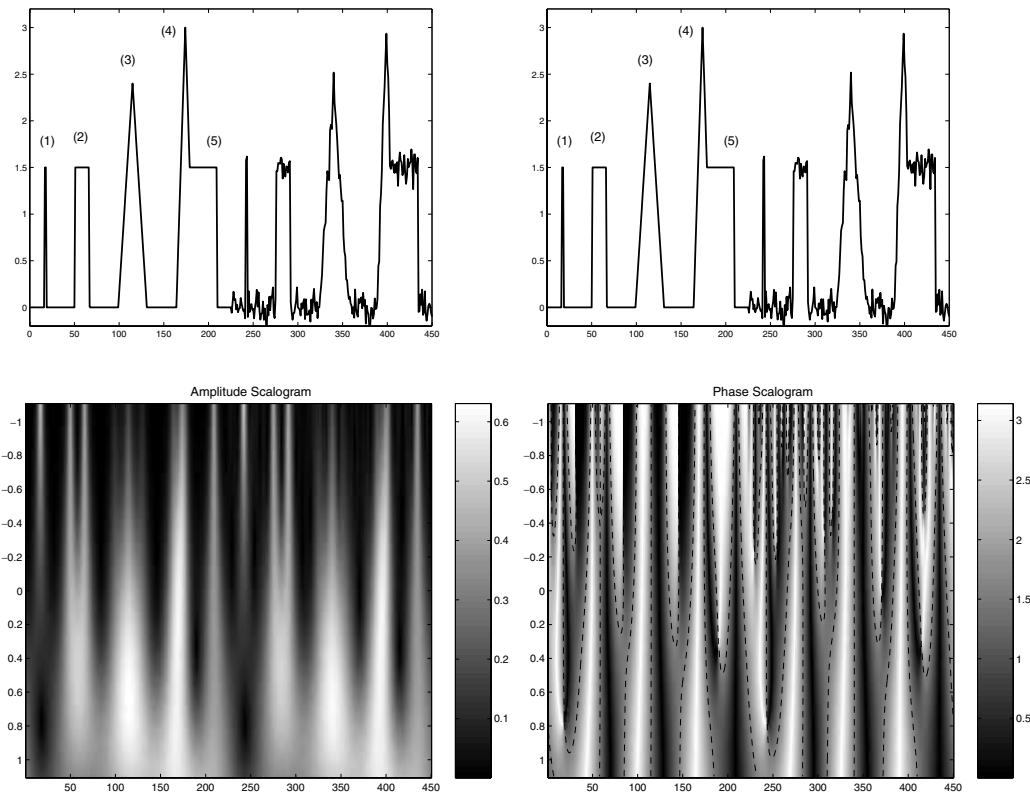


Figure 1. Example of a 1D signal and scalograms of local amplitude (left) and the absolute value of local phase (right). The level contours correspond to a local phase of $\pm \frac{\pi}{2}$ radians. A local phase of $-\frac{\pi}{2}$, $+\frac{\pi}{2}$, 0 and π corresponds to a valley, a roof(line/bar), a negative step edge and a positive step edge respectively. The scalograms are computed using a difference of Gaussian filter of 2 octaves bandwidth. The vertical axis shows the wavelength of the filter in a log scale $\ln(\frac{\lambda}{\lambda_0})$.

Phase-based feature detection has been investigated extensively following the publication of the Local Energy (LE) model of feature detection [56]. This model postulates that features can be defined and classified using their phase signatures (or their phase congruency (PC)). This observation have led to the development of a number of phase-based feature detection algorithms [18, 31, 39, 58, 61, 64–66, 74]. Phase-based feature detection (PC & LE) provides a *single* unified theory for detecting a wide range of features, rather than being specialised for a single feature type such as intensity steps.

Phase information has also been used successfully for matching and velocity estimation. For instance, phase-difference [36] and phase-correlation [44] techniques have been used for discrete two-view matching and phase gradient used for the measurement of image orientation and optical flow [23, 58].

Numerous additional advantages of the use of phase-based measurements have been reported. Perhaps the most desirable property of phase is that it is theoretically invariant to brightness and contrast. Hence it is, in principle, robust against typical variations in image

formation. However, the estimation of phase is intrinsically noisy and depends critically on the choice of the quadrature filters. It is this estimation that we investigate. The key contribution of this paper is to define quantitative performance metrics for quadrature filters to enable a meaningful comparison of filter options to be made and hence to guide the choice of filters in practice.

1.1. Local Phase and Amplitude in 1D Signals

Mathematically, the local properties (Amplitude and Phase) of a 1D real signal (function) $f(x)$ are defined using the so called analytical signal (function) $f_A(x)$ [27]:

$$f_A(x) = f(x) - i f_{\mathcal{H}}(x); \quad (1)$$

where $i = \sqrt{-1}$ and $f_{\mathcal{H}}(x)$ is the Hilbert transform of $f(x)$ defined by:

$$\begin{aligned} f_{\mathcal{H}}(x) &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{\tau - x} d\tau \\ \Leftrightarrow F_{\mathcal{H}}(\omega) &= F(\omega) \cdot i \operatorname{sign}(\omega), \end{aligned} \quad (2)$$

where, $F(\omega)$ is the Fourier transform² of $f(x)$ and

$$\text{sign}(\omega) = \begin{cases} -1 & \omega < 0 \\ +1 & \omega \geq 0 \end{cases}.$$

It is straightforward from (1) and (2) to see that:

$$F_A(\omega) = F(\omega) \cdot [1 + \text{sign}(\omega)]. \quad (3)$$

Hence, the analytical signal corresponding to f is obtained by suppressing all its negative frequencies and multiplying all its positive frequencies by two. Given the definition of the analytical signal, we can make the following definitions:

the local amplitude³ of $f(x)$:

$$A(x) = \|f_A(x)\| = \sqrt{f^2(x) + f_{\mathcal{H}}^2(x)}. \quad (4)$$

and the local phase by:

$$\phi(x) = \arg(f_A(x)) = \arctan\{f(x)/f_{\mathcal{H}}(x)\}. \quad (5)$$

From the above definition of local phase we can derive a number of phase signatures for different image features (see for e.g. [27, 74]). However, direct calculation of these local quantities, as defined above, cannot be done in a phase-based technique. This is mainly because the Hilbert transform/analytical signal, as one can see from equation (2) and (3), is defined over the whole signal/spectrum of the signal; while, for example, for feature detection, localisation in both space and frequency are highly desirable. In other words, we have to design an operator to approximate these quantities in a small, spatial span and over a narrow range of frequencies (scales) to enhance spatial localisation and to avoid the effect of noise. One way is to ‘window’ the function f , $f_w(x) = w(x) \times f(x)$, using, for example, a Gaussian function as $w(x)$ leading to the Short-Time Fourier Transform (STFT) (see for e.g. [51, chap. 4]).

An alternative approach is to estimate the local phase/energy of a filtered version of the signal. One property which this should satisfy is a zero response for a constant signal (to be invariant to grey level shift [65]), which implies that the approximation filter $f_e(x)$ must be:

- (1) a band-pass filter (zero DC).
- (2) symmetric (with constant phase, so as not to change the phase information of the original signal). The symmetry condition implies that the filter must be even (hence which is why it is denoted f_e).

We then have:

$$\begin{aligned} \hat{f}_A(x) &= f_e(x) \otimes f(x) - i \mathcal{H}(f_e(x) \otimes f(x)) \\ &= (f_e(x) - i \mathcal{H}(f_e(x))) \otimes f(x) \\ &= (f_e(x) - i f_o(x)) \otimes f(x). \end{aligned} \quad (6)$$

where “ \otimes ” denotes the 1-D convolution operator and $f_o(x)$ is the Hilbert transform of $f_e(x)$, hence they are in quadrature. It can be shown, given the definition of the Hilbert transform and the Hermitian property of the Fourier transform, that $f_o(x)$ is an odd, zero-DC band-pass filter (see e.g. [27]).

In practice, the computation of an approximation of the local phase and the local amplitude uses a pair of band-pass quadrature filters (an odd filter $f_o(x)$ and an even filter $f_e(x)$), where:

$$\hat{\phi}(x) = \arctan\{f_e(x) \otimes f(x)/f_o(x) \otimes f(x)\}, \quad (7)$$

$$\hat{A}(x) = \sqrt{[f_e(x) \otimes f(x)]^2 + [f_o(x) \otimes f(x)]^2}. \quad (8)$$

An example of an amplitude scalogram and a phase scalogram of a 1D signal is shown in Fig. 1. The signal is a combination of different type of edges; the first part of the signal is an ideal profile and the last part is a noisy version of the signal (additive white Gaussian noise with a signal-to-noise = 18 dB). Observe that all the features, independently of their shapes, are detected and can be classified using their phases. However the interpretation of a type of feature is obviously scale dependent. Consider for example the second feature: inspecting the amplitude scalogram we notice that at small scales, the feature is detected as a positive step edge followed by a negative step edge; however at very large scales there is only one peak in the amplitude scalogram and the corresponding phase at this maximum is the phase of a line. Notice for example the first feature is no longer detected at large scales; the amplitude scalogram shows a black spot and the corresponding phase is not defined, this point in scale space is called a singularity point.

1.2. Multidimensional Signals

The analytical signal and the corresponding quadrature filters provides a powerful framework for the extraction of local properties of signals. Until recently, two (and more generally N-) dimensional signal theory suffered from the absence of an *isotropic* extension of the analytical signal/Hilbert transform. Several attempts

to generalise the Hilbert transform may be found in the literature [20]: for instance, the directional/partial Hilbert transform [27, 66], the total Hilbert transform [28], and a combination of these two in the Fourier domain [28] and in the quaternionic Fourier domain [6]. Unfortunately, none of these approaches satisfy the isotropy property, which is necessary to obtain invariance with respect to orientation. Hence, other implementations have been proposed in the literature, for two-dimensional signals via the use of steerable filters [25, 72] or by performing the analysis at a set of orientations, then combining the results to provide the information at any orientation in the image [29, 31, 40, 47, 57, 58, 61, 64, 74].

Recently, Felsberg and Sommer [19, 20] proposed a novel two-dimensional generalisation of the analytic signal based on the Riesz transform, which is used instead of the Hilbert transform. A 2D isotropic analytic signal, the *monogenic signal*, is proposed. This representation preserves the core properties of the 1D analytic signal that decomposes a signal into information about structure (local phase) and energy (local amplitude). The basic idea is to design an odd isotropic filter which is vector-valued rather than scalar-valued. The “monogenic signal” is an important “phasor” for describing image properties (local phase, local amplitude and local orientation). The construction of the monogenic signal is straightforward given the 1D quadrature filter; the reader is referred to [20] for more detail.

Several applications of the monogenic signal have been already realised: edge detection and local orientation estimation [18], Moiré interferograms [45] and image denoising in a scale space framework [21].

1.3. Summary

Having introduced the analytical signal, and its extension to the multidimensional case (the monogenic signal), we have a powerful mathematical framework for the estimation of local properties of signals. These properties could be used in a number of different image processing applications. In practice, the estimation framework uses a quadrature pair of band-pass zero-DC linear filters. The construction of the multidimensional pair is straightforward given the 1D pair, or in fact given just one of the filters (the odd or the even). The question which naturally arises is: how should one choose the quadrature pair for a given application. This is the question that is addressed in the remainder of this paper.

2. Common Quadrature Filters

In this section, we study a representative set of the most common quadrature filters in the literature⁴ as well as number of new pairs derived from classical feature detectors. We provide the reader with either a mathematical expression of filter properties or a numerical estimation. For simplicity, occasionally, for a given property but different filters, we use the same symbol.

2.1. Gabor Filters

In 1946, Denis Gabor [26] introduced the windowed Fourier transform, also called the Short Time Fourier Transform (STFT). Essentially, instead of using a sinusoidal basis function $e^{i\omega x}$ which is well localised in frequency but not localised in time, he used what are now called “Gabor atoms”. These are constructed by translating in time/space and frequency a real and symmetric time/space window g :

$$g_{x_0, \omega_0}(x) = g(x - x_0) \exp[i\omega_0 x]. \quad (9)$$

The energy of g_{x_0, ω_0} is concentrated in the neighbourhood of x_0 over an interval which can be measured by the standard deviation σ_g of $\|g\|^2$. Its Fourier transform is a translation by ω_0 of the Fourier transform G of g . Therefore the energy of g_{x_0, ω_0} is concentrated near ω_0 , over an interval of size σ_G . In a time-frequency (scale-space) plane (x, ω) , the energy spread of Gabor atoms can be represented as a rectangle of width (σ_g, σ_G) centred on (x_0, ω_0) . Such a rectangle is called a “Heisenberg rectangle” [51] and the uncertainty principle shows that its area satisfies:

$$\sigma_g \sigma_G \geq \frac{1}{2}. \quad (10)$$

Equality applies only when g is a Gaussian, in which case the atoms g_{x_0, ω_0} are called Gabor functions.

Gabor functions are used extensively in digital image processing because of their optimality in the sense of the uncertainty principle, and because of their similarity to the responses of simple cells in the primary visual cortex of primates [12, 52, 59, 60]. They have been used, for instance, in texture analysis [1, 4, 15, 35, 41], texture segmentation [34, 35, 43, 75, 76], feature detection [9, 10, 57, 62], motion estimation [23, 36], image representation and coding [46], and segmentation from motion [77].

A one-dimensional Gabor function is a Gaussian modulated complex exponential:

$$g_{b(\sigma, \omega_0)}(x) = n_c \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \exp[i \omega_0 x], \quad (11)$$

$$G_{b(\sigma, \omega_0)}(\omega) = n_c \exp\left(-\frac{\sigma^2}{2}(\omega - \omega_0)^2\right), \quad (12)$$

where $\omega_0 = 2\pi/\lambda_0$ is the peak tuning frequency and n_c is a normalisation constant which is application dependent. Two alternative normalisation conditions have been used most frequently in the literature:⁵

(1) maximum condition:

$$\max_{\omega} G_b(\omega) = 1 \Rightarrow n_c = 1. \quad (13)$$

(2) constant energy condition:

$$\begin{aligned} \|g_b(x)\| &= \int_{-\infty}^{\infty} g_b^*(x) g_b(x) dx = 1 \\ \Rightarrow n_c &= \sqrt{2\sqrt{\pi} \sigma}. \end{aligned} \quad (14)$$

The half-response spatial frequency bandwidth β (in octaves) is defined as:

$$\beta = \log_2 \left(\frac{\omega_2}{\omega_1} \right), \quad (15)$$

where ω_1 and ω_2 are the solutions of $G_b(\omega)|_{n_c=1} = 1/2$ and $\omega_1 < \omega_2$. It follows that the bandwidth β of the Gabor filter(s) is the following function of $k_\beta = \sigma \omega_0$:

$$\beta(k_\beta) = \frac{\ln \left(\frac{k_\beta + \sqrt{2 \ln(2)}}{k_\beta - \sqrt{2 \ln(2)}} \right)}{\ln(2)}; \quad (16)$$

alternatively,

$$k_\beta = \sqrt{2 \ln(2)} \frac{2^\beta + 1}{2^\beta - 1}. \quad (17)$$

A complex scalar $z(\Delta x, \Delta \lambda)$ encodes the phase behaviour of the Gabor kernel, in the presence of white Gaussian noise, for a small spatial perturbation $\Delta x = x_1 - x_0$, and a small scale perturbation $\Delta \lambda = \lambda_1 - \lambda_0$,

in the neighbourhood of a scale-space point $\mathbf{p}_0(x_0 = 0, \lambda_0)$:

$$\begin{aligned} z_{g_b}(\Delta x, \Delta \lambda) &= \sqrt{\frac{2\lambda_1 \lambda_0}{\lambda_1^2 + \lambda_0^2}} \exp \left(-\frac{k_\beta^2 (\lambda_1 - \lambda_0)^2}{2(\lambda_1^2 + \lambda_0^2)} \right. \\ &\quad \left. - \frac{2\pi^2 \Delta x^2}{k_\beta^2 (\lambda_1^2 + \lambda_0^2)} + i 2\pi \frac{\lambda_1 + \lambda_0}{\lambda_1^2 + \lambda_0^2} \Delta x \right). \end{aligned} \quad (18)$$

See also Appendix A.

Unfortunately, Gabor functions have several limitations in terms of their direct application in computer vision. First, Gabor filters are non-orthogonal [27] though the family forms a frame (the family is complete but redundant) [46]. Although Gabor functions are well localised, they inevitably have infinite support. Truncation is used in most practical implementations to avoid aliasing. Finally, a Gabor pair is not in quadrature as it has non-zero negative frequencies, and a non-zero DC component (the cosine component does not integrate to zero): The normalised negative energy is:

$$\begin{aligned} \text{NE}_{g_b} &= \frac{\int_{-\infty}^0 G_b(\omega)^2 d\omega}{\int_{-\infty}^{\infty} G_b(\omega)^2 d\omega} = \frac{1}{2}(1 - \text{erf}(\sigma \omega_0)) \\ &= \frac{1}{2}(1 - \text{erf}(k_\beta)), \end{aligned} \quad (19)$$

and the DC component is:

$$\begin{aligned} \text{DC}_{g_b} &= G_b(0)|_{n_c=1} = \exp\left(-\frac{\sigma^2 \omega_0^2}{2}\right) \\ &= \exp\left(-\frac{k_\beta^2}{2}\right). \end{aligned} \quad (20)$$

Note that both the negative power and the DC component are functions of the bandwidth β . Figure 2 shows the evolution of DC_{g_b} and the root of the negative energy $\sqrt{\text{NE}_{g_b}}$ as functions of the bandwidth. We note that for a small bandwidth ($\beta < 0.7$) DC correction is not needed, as its value is close to the precision of computers (10^{-6}). However, for filters with a broad bandwidth ($\beta > 0.7$), a correction is needed. In the case of feature detection, Heitger et al. [30] and Ronse [65] independently pointed out this problem. Ronse [65] gave a second condition for the filter to be invariant to the addition of a linear ramp. Both authors proposed a DC correction for the even Gabor filter. In recent work, to which the reader is referred to for

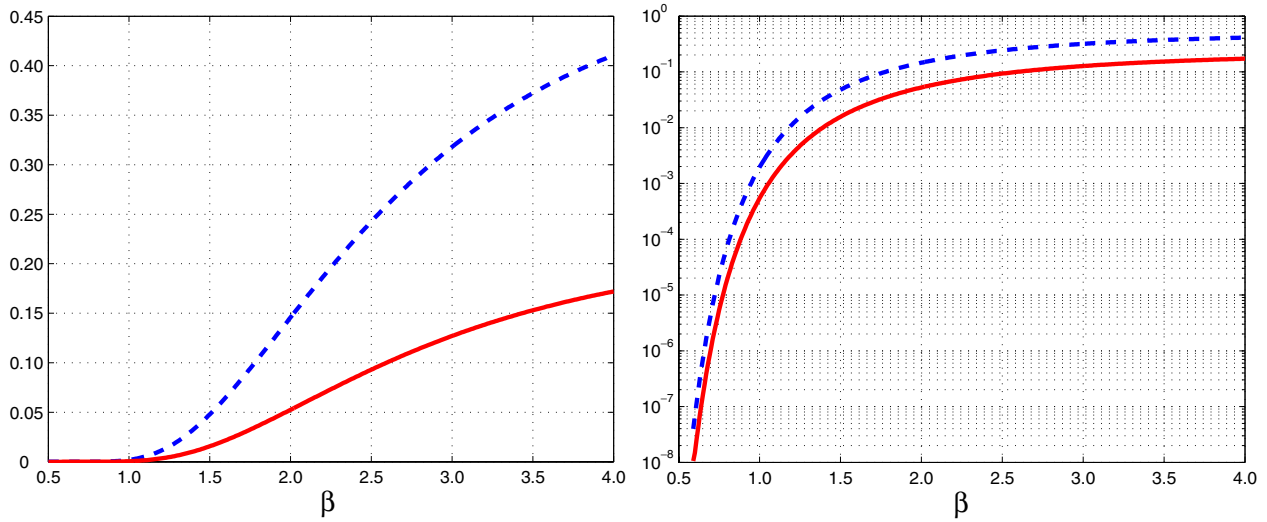


Figure 2. Evolution of DC_{gb} (dashed curve) and the negative energy $\sqrt{NE_{gb}}$ (solid curve) of Gabor filters as a function of bandwidth. Linear scale (left); logarithmic scale (right).

more detail, we studied and compared both corrections; a variant to Ronse's correction is also proposed [3].

2.2. Log-Gabor Filters

Log-Gabor (log-Normal) filters have become the preferred choice in many computer vision applications using quadrature filters as they allow arbitrarily large bandwidth zero DC filters to be constructed [18, 19, 22, 27, 40, 58]. These filters are particularly suited for the estimation of the local frequency using the ratio of log-Gabor filters [27].

The one-dimensional log-Gabor filter in the frequency domain is a Gaussian function on a logarithmic scale:

$$G_I(\omega) = n_c \exp\left(-\frac{\ln^2(\omega/\omega_0)}{2 \ln^2(\kappa_\beta)}\right), \quad (21)$$

where ω_0 is the peak tuning frequency and $0 < \kappa_\beta < 1$ is related to the bandwidth β of the filter by the following relation:

$$\beta = -\frac{2\sqrt{2}}{\sqrt{\ln 2}} \ln(\kappa_\beta), \quad (22)$$

and inversely:

$$\kappa_\beta = \exp\left(-\frac{1}{4} \sqrt{2 \ln(2)} \beta\right). \quad (23)$$

The two values for the normalisation constant n_c (defined by Eqs. (13) and (14)) are given by:

$$\max_{\omega} G_I(\omega) = 1 \Rightarrow n_c = 1, \quad (24)$$

$$\begin{aligned} \|g_I(x)\| &= \frac{1}{2\pi} \|G_I(\omega)\| = 1 \\ \Rightarrow n_c &= \exp\left(-\frac{1}{8} \ln(\kappa_\beta)^2\right) \sqrt{-\frac{2\sqrt{\pi}}{\omega_0 \ln(\kappa_\beta)}}. \end{aligned} \quad (25)$$

The main advantage of log-Gabor filters is that as well as being invariant to a grey-level shift (zero DC) they are also invariant to an additive ramp.⁶ A linear operator $f(x)$ is invariant to an additive ramp if and only if:

$$\int_{-\infty}^{\infty} x f(x) dx = 0. \quad (26)$$

Using the moment property of the Fourier transform, we have:

$$\int_{-\infty}^{\infty} x g_I(x) dx = i \frac{d}{d\omega} G_I(\omega)|_{\omega=0} = 0. \quad (27)$$

Care must be taken in the design of log-Gabor filters to avoid aliasing. Indeed, the long tails of the filters may run beyond the Nyquist limit imposed by the sampling frequency. Hence, we would like the Nyquist limit to be at least n times the standard deviation away from ω_0 on a logarithmic scale. This gives the following

condition:

$$\omega_0 < \pi \kappa_\beta^n. \quad (28)$$

In practice, $n = 3$ is usually sufficient [27].

In addition to the above limit for the central frequency, probably the most significant drawback is the absence of an analytical expression for the filter in the spatial domain.

2.3. Gaussian Derivative Filters

Gaussian derivatives are well known in computer vision. This is because of some very desirable properties of the Gaussian function (used as the smoothing function), such as simplicity and steerability. For instance, Gaussian derivatives have been used in feature detection [8, 17, 25, 37, 47, 61, 65, 66] and motion estimation [33, 49].

It is well known that the Fourier transform of a Gaussian function is a Gaussian, and using the derivative property of the Fourier transform, we define the following family of Gaussian derivative quadrature filters in the frequency domain:

$$G_d(\omega) = \begin{cases} n_c \omega^a \exp[-\sigma^2 \omega^2], & \text{if } \omega \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (29)$$

where, $a \in \mathbb{R}$ and $a \geq 1$.

Here, we relax the constraint on a and hence we do not have an analytical expression of the filters in the temporal/spatial domain for $a \notin \mathbb{N}$.

The Hilbert transform of the Gaussian function is the following function (e.g. see [65, 66]):

$$g_{\mathcal{H}}(x) = \frac{\sqrt{2}}{\pi \sigma} \text{Dw} \left[\frac{x}{\sqrt{2} \sigma} \right], \quad (30)$$

where

$$\text{Dw}(x) = \exp(-x^2) \int_0^x \exp(t^2) dt$$

is the *Dawson's integral* function (31)

and

$$\frac{d}{dx} \text{Dw}(x) = 1 - 2x \text{Dw}(x). \quad (32)$$

The *Dawson's integral* function and its derivatives are shown in Fig. 3. When $a \in \mathbb{N} \subset \mathbb{R}$ and odd, the corresponding odd/even filter in the spatial domain is the a th

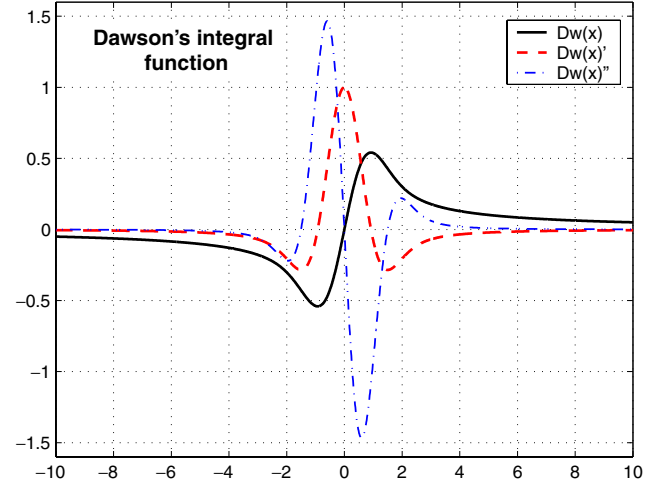


Figure 3. The *Dawson's integral* function and its derivatives in the spatial domain.

derivative of the Gaussian/ $g_{\mathcal{H}}(x)$, and for even values of a the corresponding odd/even filter is the a th derivative of $g_{\mathcal{H}}(x)$ /Gaussian function. This is up to a change of sign and a normalisation factor of $n_c/2$.

It is straightforward to see that the family has zero DC, and that it is therefore invariant to a grey level shift, and for $a > 1$, it is invariant to an additive ramp (see (26)). The two values for the normalisation constant n_c are given by:

$$\max_{\omega} G_d(\omega) = 1 \Rightarrow n_c = 1/G_d(\omega_0), \quad (33)$$

$$\begin{aligned} \|g_d(x)\| &= \frac{1}{2\pi} \|G_d(\omega)\| = 1 \\ \Rightarrow n_c &= 2 \frac{\sqrt{\pi} \sigma^{(a+\frac{1}{2})}}{\sqrt{\Gamma(a+\frac{1}{2})}}, \end{aligned} \quad (34)$$

where $\Gamma(\cdot)$ is the gamma function and ω_0 is the peak tuning frequency given by:

$$\omega_0 = \frac{\sqrt{a}}{\sigma}. \quad (35)$$

The octave bandwidth of this family is given by:

$$\beta = \frac{\ln \left(\frac{W(-1, c)}{W(0, c)} \right)}{2 \ln(2)}, \quad \text{with } c = -\frac{4^{-(1/a)}}{e}. \quad (36)$$

Here, $W(k, \cdot)$ is the k th branch of the Lambert's W function [11].

The Gaussian derivative kernels, as defined here, have the advantage of allowing a continuous bandwidth

with a maximum bandwidth of 2.59 octaves corresponding to the 1st derivative. Note that the analytical expression in the spatial domain is only valid for integer derivative values. Differential Gaussian filters exhibit less aliasing than log-Gabor filters.

2.4. Difference of Gaussian Filter (DoG)

The DoG filter has been used in a number of feature detection applications [42, 48, 50, 53, 55] and has been proposed as a filtering model of the human visual system. For example, retinal ganglion cells and lateral geniculate cells early in the visual processing have receptive fields which can be implemented as DoG filters [53, and references therein].

The filter (even) is given by:

$$f_{\text{DoGe}}(x) = \frac{n_c}{\sqrt{2\pi} \sigma'} \exp\left(-\frac{x^2}{2\sigma'^2}\right) - \frac{n_c}{\sqrt{2\pi} \sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (37)$$

where $\sigma > \sigma' > 0$.

This operator approximates well the second derivative of a Gaussian and has one more degree of freedom.⁷ Several authors have specified a preference for it possibly because of its ease of implementation. However, this filter has a far better performance, for line detection, than the 2nd derivative of a Gaussian in the sense of Canny's criteria.

As already noted, the Hilbert transform of the Gaussian function is the Dawson's function given by (30). Hence the odd filter in quadrature with $f_{\text{DoGe}}(x)$ is given by the following:

$$f_{\text{DoGo}}(x) = n_c \frac{\sqrt{2}}{\pi \sigma'} \text{Dw}\left[\frac{x}{\sqrt{2} \sigma'}\right] - n_c \frac{\sqrt{2}}{\pi \sigma} \text{Dw}\left[\frac{x}{\sqrt{2} \sigma}\right]. \quad (38)$$

Therefore, the Fourier transform of the quadrature pair is defined by (up to a multiplicative factor of 2):

$$F_{\text{DoG}}(\omega) = \begin{cases} n_c \exp\left(-\frac{\sigma'^2 \omega^2}{2}\right) - n_c \exp\left(-\frac{\sigma^2 \omega^2}{2}\right), & \text{if } \omega \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (39)$$

and the peak tuning frequency is given by:

$$\omega_0 = \frac{2}{\sigma} \left(\frac{\ln(\gamma)}{(\gamma^2 - 1)} \right)^{1/2} \quad \text{where, } \gamma = \frac{\sigma'}{\sigma}. \quad (40)$$

The two values for the normalisation constant n_c for this family are given by:

$$\max_{\omega} F_{\text{DoG}}(\omega) = 1 \Rightarrow n_c = 1/F_{\text{DoG}}(\omega_0), \quad (41)$$

$$\begin{aligned} \|f_{\text{DoG}}(x)\| &= \frac{1}{2\pi} \|F_{\text{DoG}}(\omega)\| = 1 \\ \Rightarrow n_c &= 2\sqrt{\sqrt{\pi} \sigma} \left(\frac{1}{\gamma} - \frac{2\sqrt{2}}{\sqrt{1+\gamma^2}} + 1 \right)^{-\frac{1}{2}}. \end{aligned} \quad (42)$$

The quadrature DoG filter is invariant to grey level shift (zero DC) and invariant to an additive ramp (use Eq. (27) with Eq. (39)).

Unfortunately, we have not been able to derive an analytical expression for the bandwidth of the filter. It's straightforward to note that the bandwidth depends only on the ratio $\gamma = \sigma'/\sigma$. The bandwidth as a function of γ is shown in Fig. 4 (a). Notice that the smallest bandwidth allowable for this family is 1.76 (corresponding to the bandwidth of the 2nd derivative of a Gaussian). The filters for $\lambda_0 = 20$ and for different values of γ are shown in Fig. 4 (b) and (c) on a linear and logarithmic scale respectively. Observe that the tuning frequency must be chosen carefully to avoid aliasing problem especially for small values of γ .

2.5. Cauchy Filter

Cauchy filters have been proposed in [38, 65, 66]. This family does not appear to be well known in the computer vision community. However, this family has some attractive properties as we see below. The Cauchy functions are given by:

$$f_n(x) = \Re\{(1 - ix)^{-n-1}\} = \frac{\Re\{(1 + ix)^{n+1}\}}{(1 + x^2)^{n+1}}; \quad (43)$$

$$g_n(x) = \Im\{(1 - ix)^{-n-1}\} = \frac{\Im\{(1 + ix)^{n+1}\}}{(1 + x^2)^{n+1}}; \quad (44)$$

where $n \in \mathbb{N}$ and determines the number of ripples of the two functions and $\Re(z)/\Im(z)$ is the real/imaginary part of z . f_n and g_n are in quadrature and are integrable

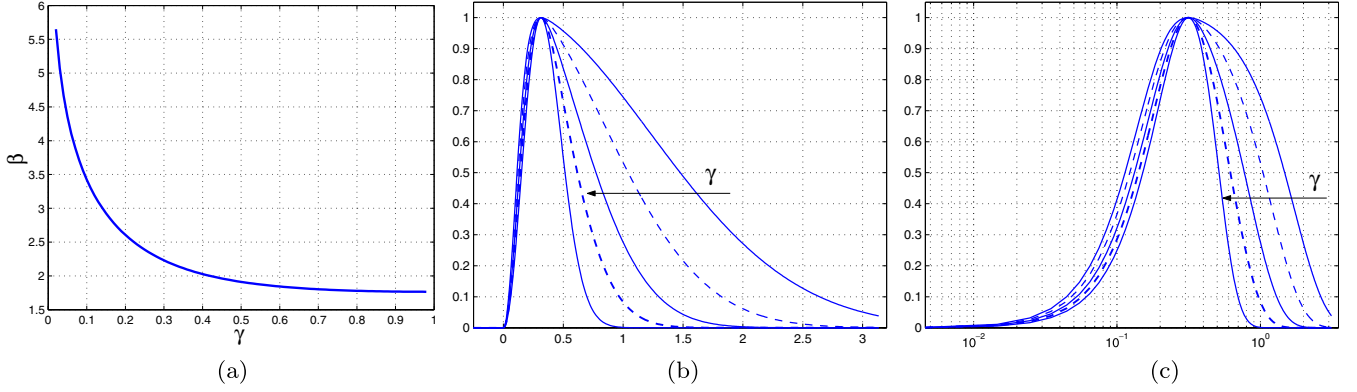


Figure 4. Difference of Gaussian filter. Bandwidth as a function of γ (a); Fourier transform of the filter on a linear (b) and logarithmic (c) scale for $\lambda_0 = 20$ and $(\gamma, \beta) = (0.081, 3.63), (0.131, 3.05), (0.212, 2.53), (0.343, 2.12)$ and $(0.9, 1.76)$. When $\gamma \approx 0.9$ the filter is a good approximation to the 2nd derivative of a Gaussian.

for $n \geq 1$. The Fourier transform of f_n is given by [66]:

$$F_n(\omega) = 2\pi(n!)^{-1}|\omega|^n \exp(-|\omega|). \quad (45)$$

The values $n = 3$ and $n = 5$ have been proposed to model some properties of the human visual response in [38].

Following the same procedure as for the Gaussian derivatives, we can define a family of quadrature filters using Cauchy functions for $(n = a \geq 1) \in \mathbb{R}$; which in the Fourier domain gives:

$$F_{Ch}(\omega) = \begin{cases} n_c \omega^a \exp(-\sigma \omega) & \text{if } \omega \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (46)$$

where σ is a scaling parameter. Hence, by definition, the filters have zero DC and are invariant to an additive ramp for $a > 1$. Notice that the expression of the filters is similar to that for the Gaussian derivative filters. It can be shown that the filters in the spatial domain are given by:

$$\begin{aligned} f_{Che}(x) &= n_c \frac{\Gamma(a+1)}{2\pi \Re(\sigma - ix)^{a+1}} \\ &= n_c \frac{\Gamma(a+1)}{2\pi} \frac{\Re(\sigma + ix)^{a+1}}{(\sigma^2 + x^2)^{a+1}}, \end{aligned} \quad (47)$$

$$\begin{aligned} f_{Cho}(x) &= n_c \frac{\Gamma(a+1)}{2\pi \Im(\sigma - ix)^{a+1}} \\ &= n_c \frac{\Gamma(a+1)}{2\pi} \frac{\Im(\sigma + ix)^{a+1}}{(\sigma^2 + x^2)^{a+1}}, \end{aligned} \quad (48)$$

and the normalisation constants are given by:

$$\max_{\omega} F_{Ch}(\omega) = 1 \Rightarrow n_c = 1/F_{Ch}(\omega_0), \quad (49)$$

$$\begin{aligned} \|f_{Ch}(x)\| &= \frac{1}{2\pi} \|F_{Ch}(\omega)\| = 1 \\ \Rightarrow n_c &= \left(\frac{\pi 4^{(a+1)} \sigma^{(2a+1)}}{\Gamma(2a+1)} \right)^{\frac{1}{2}}, \end{aligned} \quad (50)$$

where ω_0 is the peak tuning frequency of the filter and is given by the following ratio:

$$\omega_0 = \frac{a}{\sigma}. \quad (51)$$

The octave bandwidth of this family is given by:

$$\beta = \frac{\ln\left(\frac{W(-1,c)}{W(0,c)}\right)}{\ln(2)}, \quad \text{with } c = -\frac{2^{-(1/a)}}{e}. \quad (52)$$

Finally, the complex scalar $z(\Delta x, \Delta \lambda)$ which encodes the phase behaviour of this family for a small spatial perturbation Δx , and a small scale perturbation $\Delta \lambda$ is given by:

$$\begin{aligned} z_{Ch}(\Delta x, \Delta \lambda) &= \left(\frac{2\sqrt{\sigma_1 \sigma_0}}{\sigma_1 + \sigma_0 - i\Delta x} \right)^{2a+1} \\ &= \left(\frac{2a\sqrt{\lambda_1 \lambda_0}}{a(\lambda_1 + \lambda_0) - i2\pi \Delta x} \right)^{2a+1}. \end{aligned} \quad (53)$$

Figure 5(a) shows the evolution of the bandwidth of the filters function of a . The bandwidth of the Gaussian derivatives is shown for comparison. The transfer function of the quadrature pair is shown on a linear and

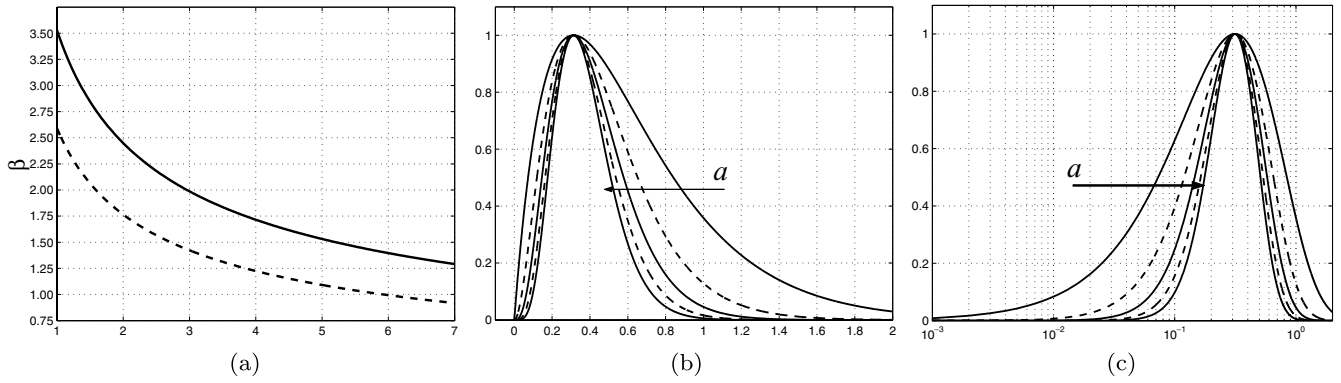


Figure 5. Cauchy quadrature filters; Bandwidth β as a function of a ; Cauchy filters (solid curve) and Gaussian derivative filter (dashed curve) (a); Fourier transform of the quadrature pair on a linear (b) and logarithmic (c) scale for $a = 1, 2, 3, 4$ and 5 .

logarithmic scale in Fig. 5(b) and (c). Note the similarity of Cauchy functions to the Rayleigh distribution. This may suggest their use as a computationally fast approximation to Rayleigh.

2.6. Quadrature Filters Related to Canny's Criteria

Several papers have considered designing optimal edge/line detectors in the sense of Canny's criteria [8] (see Appendix B). These detectors are typically the first/2nd derivative of a smoothing function. To our knowledge, two smoothing functions have been proposed in this sense.

2.6.1. Deriche Filter. We can define a family of quadrature filters using the smoothing function proposed by Deriche [13]. This function is the integral of the optimal operator in the sense of Canny's criteria for edge detection [8]. It is also the smoothing function used by Ziou [80] for optimal line detection according to Canny's criteria. Hence, it is the second integral of the optimal line detector and is given by:

$$s_d(x) = (\alpha|x| + 1) \exp(-\alpha|x|), \quad (54)$$

$$S_d(\omega) = \frac{4\alpha^3}{(\omega^2 + \alpha^2)^2}, \quad (55)$$

where α is the scaling parameter ($\sim \sigma^{-1}$).

In a similar way as for Gaussian derivative filters we define a family of quadrature filters in the Fourier

domain:

$$F_d(\omega) = \begin{cases} n_c \omega^a S_d(\omega), & \text{if } \omega \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad \text{where, } 1 \leq a \leq 2. \quad (56)$$

The odd filter is the Deriche edge detector for $a = 1$ and the even filter is Ziou's line detector when $a = 2$. The filters are invariant to a grey level shift, and invariant to an additive ramp for $a > 1$ (see Eq. (26)).

It can be shown that the peak tuning frequency is the following function:

$$\omega_0 = \alpha \sqrt{\frac{a}{4-a}}, \quad (57)$$

and, the two values for the normalisation constant n_c are given by:

$$\max_{\omega} F_d(\omega) = 1 \Rightarrow n_c = 1/F_d(\omega_0), \quad (58)$$

$$\begin{aligned} \|f_d(x)\| &= \frac{1}{2\pi} \|F_d(\omega)\| = 1 \\ \Rightarrow n_c &= \alpha^{(\frac{1}{2}-a)} \sqrt{\frac{-12 \cos(\pi a)}{(2a-5)(2a-3)(2a-1)}}. \end{aligned} \quad (59)$$

Unfortunately, there is not a simple analytical expression for the bandwidth of the filter. We did a numerical estimation of the bandwidth and approximated its functional dependence on a by the following 4th order

Table 1. Coefficients of the fitting function given by Eq. (60).

i	c_i	μ_i
0	2.54372940760	–
1	0.11940325096	2.00000000000
2	0.01269121401	0.58167860542

polynomial:

$$\beta = c_0 + c_1 \left(\frac{a - \mu_1}{\mu_2} \right)^2 + c_2 \left(\frac{a - \mu_1}{\mu_2} \right)^4. \quad (60)$$

The coefficients are tabulated in Table 1. The maximum relative error with respect the numerical estimation is 0.0006.

2.6.2. Sarkar and Boyer Filter. Sarkar and Boyer [67] proposed a modification of the multiple response criteria appropriated to the use of IIR filters. Following a similar approach as Canny [8] and Deriche [13], the authors proposed the following smoothing function:

$$s_s(x) = \frac{1}{\alpha} \exp(-\alpha |x|) \times \left\{ \cos(\varphi) - \frac{\cos(\alpha \tilde{m} |x| + \varphi) - \tilde{m} \sin(\alpha \tilde{m} |x| + \varphi)}{\tilde{m}^2 + 1} \right\}, \quad (61)$$

where φ and \tilde{m} determine the shape of the filter and α is the scaling parameter. Note that the above smoothing function is equivalent to Eq. (54) for $\varphi = \frac{\pi}{2}$ and when $\tilde{m} \rightarrow 0$.

The Fourier transform of (61) is:

$$S_s(\omega) = 2\alpha^2 \tilde{m} \times \frac{(2 \sin(\varphi) - \tilde{m} \cos(\varphi))\omega^2 + \alpha^2 (\tilde{m}^3 \cos(\varphi) + 3\tilde{m} \cos(\varphi) + 2 \sin(\varphi))}{(\omega^2 + \alpha^2)((\omega + \alpha \tilde{m})^2 + \alpha^2)((\omega - \alpha \tilde{m})^2 + \alpha^2)}, \quad (62)$$

Hence, we can define the following family of quadrature filters in the Fourier domain:

$$F_s(\omega) = \begin{cases} n_c \omega^a S_s(\omega), & \text{if } \omega \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad \text{where, } 1 \leq a \leq 2. \quad (63)$$

Unfortunately, the Fourier representation of the filters is not easy to manipulate analytically. Therefore, we studied the filters numerically for $a \in [1, 2]$, and for 4 combinations of the shape parameters (\tilde{m}, φ) corresponding to the best values of $\Omega_S(f_s)$, $\Sigma\Gamma(f_s)$, $\Sigma\Gamma = \Omega_S(f_s)$ and $\Sigma\Gamma\Omega_S(f_s)$ for the odd part when $a = 1$ (i.e Sarkar's edge operator). The values are tabulated in Table 2. Figure 6(a) shows the evolution of the bandwidths as a function of a . We first notice a high correlation between the SNR-localisation and the bandwidth

Table 2. Bandwidth and peak tuning frequency of $F_s(\omega)$ for some particular values of the shape parameters.

	$\Sigma\Gamma$	Ω_S	\tilde{m}	φ	ω_0/α	β	
$a = 1$	0.885	3.252	1.917	0.651	1.354	2.7007	(*)
	2.000	1.622	0.002	1.382	0.5772	3.0096	
	1.833	1.833	0.074	0.120	0.4916	2.9332	
	1.443	2.453	0.464	0.332	0.5355	2.7875	
GD	0.921	3.245	–	–	$\omega_0 = 1/\sigma$	2.5902	
$a = 2$	–	–	1.917	0.651	1.7468	1.5554	(*)
	–	–	0.002	1.382	1.0000	2.5431	
	–	–	0.074	0.120	0.8147	2.4012	
	–	–	0.464	0.332	0.8237	2.1393	
GD	–	–	–	–	$\omega_0 = \sqrt{2}/\sigma$	1.7647	

$\Sigma\Gamma$ and Ω_S are given when the odd filter is used for edge detection alone. The Gaussian derivative filter is shown for comparison.

(*) Here the filter is equivalent to the Deriche filter, $\tilde{m} = 0.002$ is the smallest value in our sampling space. Basically the true values for the Deriche filters should be $\tilde{m} = 0$ and $\varphi = \frac{\pi}{2}$.

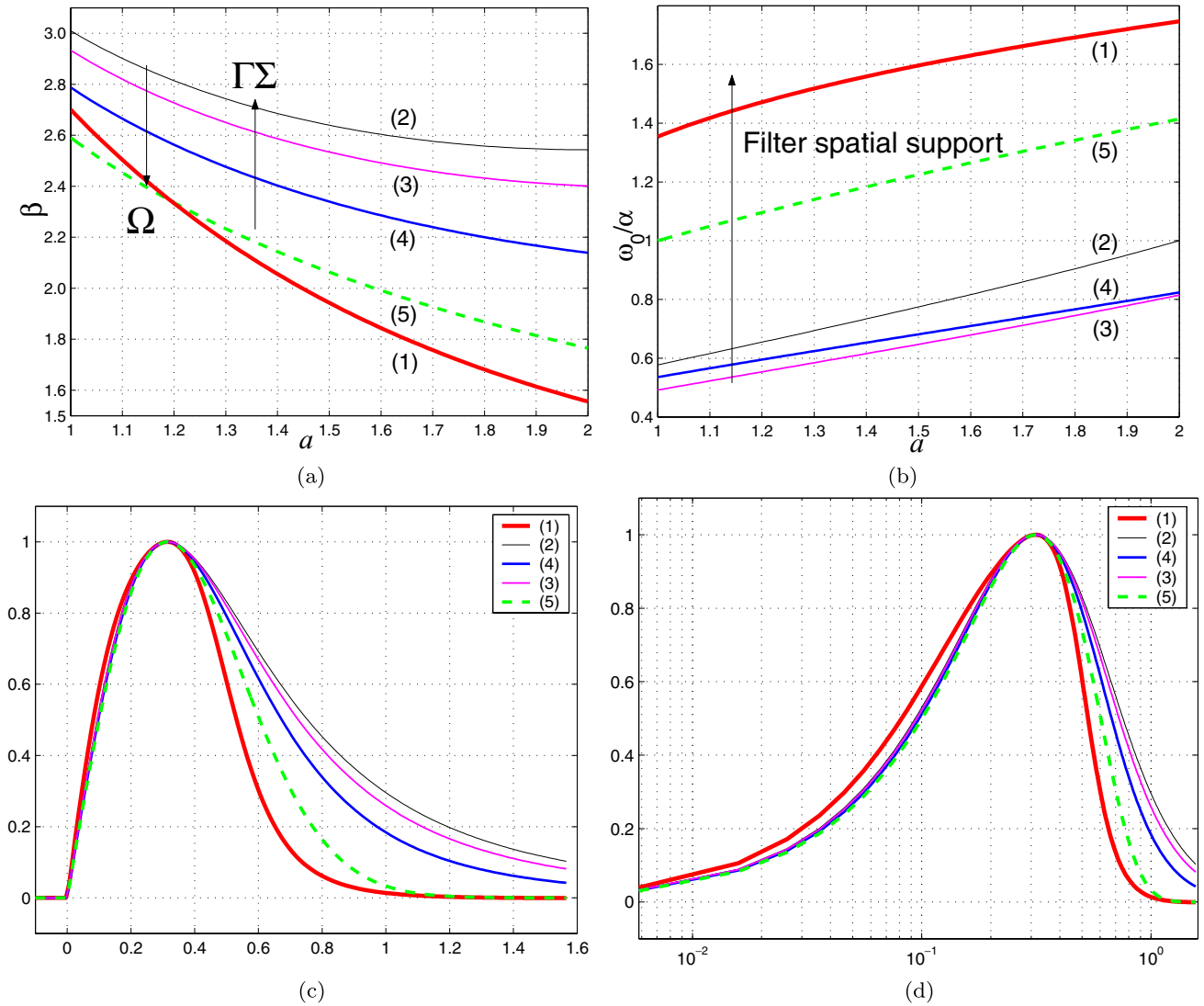


Figure 6. Sarkar and the Deriche filters. Bandwidth as a function of a (a); ω_0/α as a function of a (b); Filters in the Fourier domain on a linear (c) and logarithmic (d) scale for $\lambda_0 = 20$ and $a = 1$. best Ω_S (Red thick curves (1)); best $\Sigma\Gamma$ (Black thin curves (2)); $\Omega_S = \Sigma\Gamma$ (Magenta curves (3)); best $\Sigma\Gamma\Omega_S$ (Blue curves (4)) and Gaussian derivative filter (Green dashed curves (5)). The numbers refer to the same order in Table 2.

of the filters; $\Sigma\Gamma$ increases when β increases. The 1st derivative of the Gaussian is an exception to this rule. We believe that this is mainly due to the difference in the decay of the Gaussian function (exponential) and Sarkar's smoothing function (ω^{a-4}). This is illustrated in Fig. 6(c) and (d). Values of ω_0/α give some insight regarding the spatial support of the filters (see Fig. 6(b)).

We might expect that the performance of a detector depends on its spatial shape. However, this experiment shows the importance of the bandwidth of the filter. Notice that the smallest bandwidth for the edge operators is 2.6 octaves. Several authors mention that larger bandwidths give better results in phase based feature

detection [31, 40, 58]. A bandwidth of 1.5 octaves is used in [30, 31, 39, 64], 1.76 octaves in [47, 61] and 2 octaves in [40, 58]; all smaller than 2.6 and even smaller than 2.54 (bandwidth of Ziou's line detector). The second important thing that we can learn from these values, is that even the smallest bandwidth for the edge operators is larger than the largest one for line detectors. Knowing for example that, for the Deriche filter, the odd part alone, for $a = 1$, is optimal in the sense of Canny for edge detection and the even part, for $a = 2$, is optimal in the same sense for lines,⁸ this implies that a quadrature pair will not be optimal for both in the sense of Canny. Hence, a tradeoff must be made.

The spatial representation of the filters is shown in Fig. 7. Figure 7(a) and (b) shows the filters for $a = 1$ (i.e odd parts correspond to the edge operators) and Fig. 7(c) and (d) shows the same filters but for a constant bandwidth of 2.5431 octaves (i.e different values of a). Interestingly, the shape (by which we mean the relative differences) of the filters are similar for both cases. Hence we expect the filters to have similar relative performances in the sense of Canny. This experiment shows the importance of filter shape. Notice that a bandwidth of 2.4531 corresponds to $a = 2$ for the Deriche filter (optimal even part) and to $a = 1.032642$ for the Gaussian derivatives (good approximation to the 1st derivative). Inspecting Fig. 7(d), it is more likely that the odd part of a Deriche filter for $a = 2$

will have a better $\Sigma\Gamma$ than the 1st derivative of a Gaussian.

The above experiments highlight issues in choosing a quadrature pair for feature detection and on which basis. It also suggests the possibility of adapting Canny's criteria to the local energy model for the purpose of filter design aiming to have similar performances at least for edges and lines. This is a very delicate point since, for example, the localisation criteria and the multiple responses criteria might not be suitable as they do not take into account the phase information. However, the adaptation of the SNR criteria is straightforward.

We do not have a clear answer to the above questions but they are worthy of further investigation. It is more likely and justified, however, to use a quadrature pair

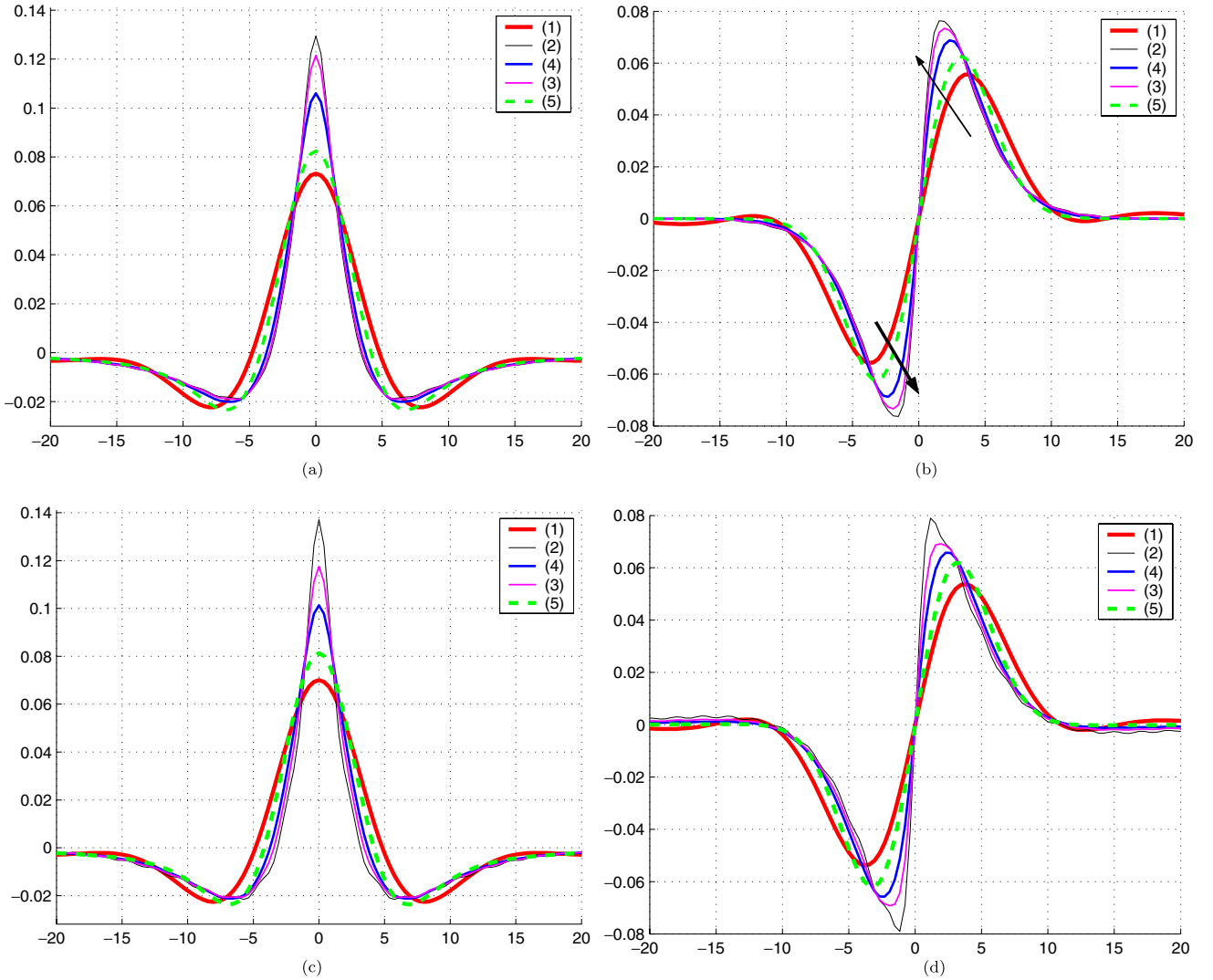


Figure 7. Sarkar and the Deriche filters in the spatial domain for $\lambda_0 = 20$. Constant value of $a = 1$ (a) and (b), the arrows show increasing $\Sigma\Gamma$ (see Fig. 6 (c)–(d) for the Fourier representation); Constant bandwidth $\beta = 2.5431$ octaves (c) and (d). Best Ω_S (Red thick curves); best $\Sigma\Gamma$ (Black thin curves); $\Omega_S = \Sigma\Gamma$ (Magenta curves); best $\Sigma\Gamma\Omega_S$ (Blue curves) and Gaussian derivative filter (Green dashed curves).

where the odd/even part is optimal in the sense of Canny if we are more interested in detecting edges/lines, assuming that the contribution of the even/odd part of the filter is negligible.

3. Comparison

We have presented most of the mathematical properties of the different families in the previous section. A summary of the key equations and properties is given in Table 3. In this section, we present a comparison of the filters used as a quadrature pair. For this purpose, the comparison will be fair only if the filters have the same bandwidth and the same tuning frequency.

Figure 8 shows the five different families of filters (Deriche, DoG, Gaussian Derivative, Cauchy and log-Gabor) for a bandwidth of 2.5431 octaves (Bandwidth of Ziou's line detector) and a wavelength of 20, both in the frequency domain and in the spatial domain. Inspecting the Fourier representation of the filters, we notice that the Gaussian derivative filters have the least aliasing, and the Deriche filter is the worst. Notice also that the three families, Deriche, log-Gabor and the DoG filter are very similar for this bandwidth, and the Cauchy family is between these three filters and the Gaussian derivative filter. This gives us some insight into the behaviour of the filters for feature detection as we know that the even part of the Deriche filter is optimal in the sense of Canny, and the odd part of the Gaussian has a high Ω_S and a low $\Sigma\Gamma$. The log-

Gabor filter is an acceptable approximation of Ziou's line detector (even part of the Deriche filter), and has low aliasing. Hence, if we use a Fourier domain implementation, this suggests that the log-Gabor filter is preferable for line detection using this bandwidth.

Figure 9 shows the same five families for the same wavelength of 20 but for a larger bandwidth ($\beta = 3.0096$ octaves) in the frequency domain and in the spatial domain, except for the Gaussian derivative filter where $\beta = 2.5902$ (maximum bandwidth corresponding to the 1st derivative). A bandwidth of 3.0096 octaves corresponds to the Deriche edge operator for the Deriche family which has the best SNR-localisation. First of all, notice that for this bandwidth, the Deriche filter has less aliasing than the log-Gabor filter. This is the opposite relation to before (i.e, when $\beta = 2.5431$ octaves). We remark that the Cauchy filter approximates the Deriche filter quite well and has the advantage of being invariant to an additive ramp besides the fact that Cauchy filters has less aliasing than the Deriche filter. The DoG filter seems to be very similar to the log-Gabor filter except that the DoG filter has less aliasing. Notice that, the DoG filter with this bandwidth (i.e $\gamma = 0.1350$) has a better SNR-localisation than Ziou's optimal line detector (even part of the Deriche filter for $a = 2$) and similar performance in term of the product of the three criteria.

3.1. Phase Stability

Phase stability is very important in phase-based estimation techniques. Stability over scales is important

Table 3. Summary of main equations and filters properties.

	Temporal/Spatial domain	Fourier domain	Bandwidth	Tuning frequency	Zero-DC/ramp invariance
Gabor	$n_c \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} e^{i\omega_0 x}$	$n_c e^{-\frac{\sigma^2}{2}(\omega-\omega_0)^2}$	$\frac{\ln\left(\frac{\sigma\omega_0+\sqrt{2\ln(2)}}{\sigma\omega_0-\sqrt{2\ln(2)}}\right)}{\ln(2)}$	ω_0	No/No
log-Gabor	No	$n_c \exp\left(-\frac{\ln^2(\omega/\omega_0)}{2\ln^2(\kappa\beta)}\right)$	$-\frac{2\sqrt{2}}{\sqrt{\ln 2}} \ln(\kappa\beta)$	ω_0	Yes/Yes
Gaussian derivative	Only for $a \in \mathbb{N}$	$n_c \omega^a \exp[-\sigma^2\omega^2]$	$\frac{\ln\left(\frac{W(-1,c)}{W(0,c)}\right)}{2\ln(2)}, c = -\frac{4^{-(1/a)}}{e}$	$\frac{\sqrt{a}}{\sigma}$	Yes/Yes ($a > 1$)
DoG	Yes	$n_c \left(e^{-\frac{\sigma'^2\omega^2}{2}} - e^{-\frac{\sigma^2\omega^2}{2}}\right)$	Numerical (Fig. 4 (a))	$\frac{2}{\sigma} \sqrt{\frac{\ln(\gamma)}{\gamma^2-1}}, \gamma = \frac{\sigma'}{\sigma}$	Yes/Yes
Cauchy	$n_c \frac{\Gamma(a+1)}{2\pi(\sigma-ix)^{a+1}}$	$n_c \omega^a \exp(-\sigma\omega)$	$\frac{\ln\left(\frac{W(-1,c)}{W(0,c)}\right)}{\ln(2)} c = -\frac{2^{-(1/a)}}{e}$	$\frac{a}{\sigma}$	Yes/Yes ($a > 1$)
Deriche	No	$n_c \omega^a \frac{4\alpha^3}{(\omega^2+\alpha^2)^2}$	Numerical Fig. 6(a)	$\alpha \sqrt{\frac{a}{4-a}}$	Yes/Yes ($a > 1$)
Sarkar and Boyer	No	Yes (*)	Numerical Fig. 6 (a)	Numerical Fig. 6 (b)	Yes/Yes ($a > 1$)

$$(*) 2 n_c \omega^a \alpha^2 \tilde{m} \frac{(2 \sin(\varphi) - \tilde{m} \cos(\varphi)) \omega^2 + \alpha^2 (\tilde{m}^3 \cos(\varphi) + 3 \tilde{m} \cos(\varphi) + 2 \sin(\varphi))}{(\omega^2 + \alpha^2)((\omega + \alpha \tilde{m})^2 + \alpha^2)((\omega - \alpha \tilde{m})^2 + \alpha^2)}.$$

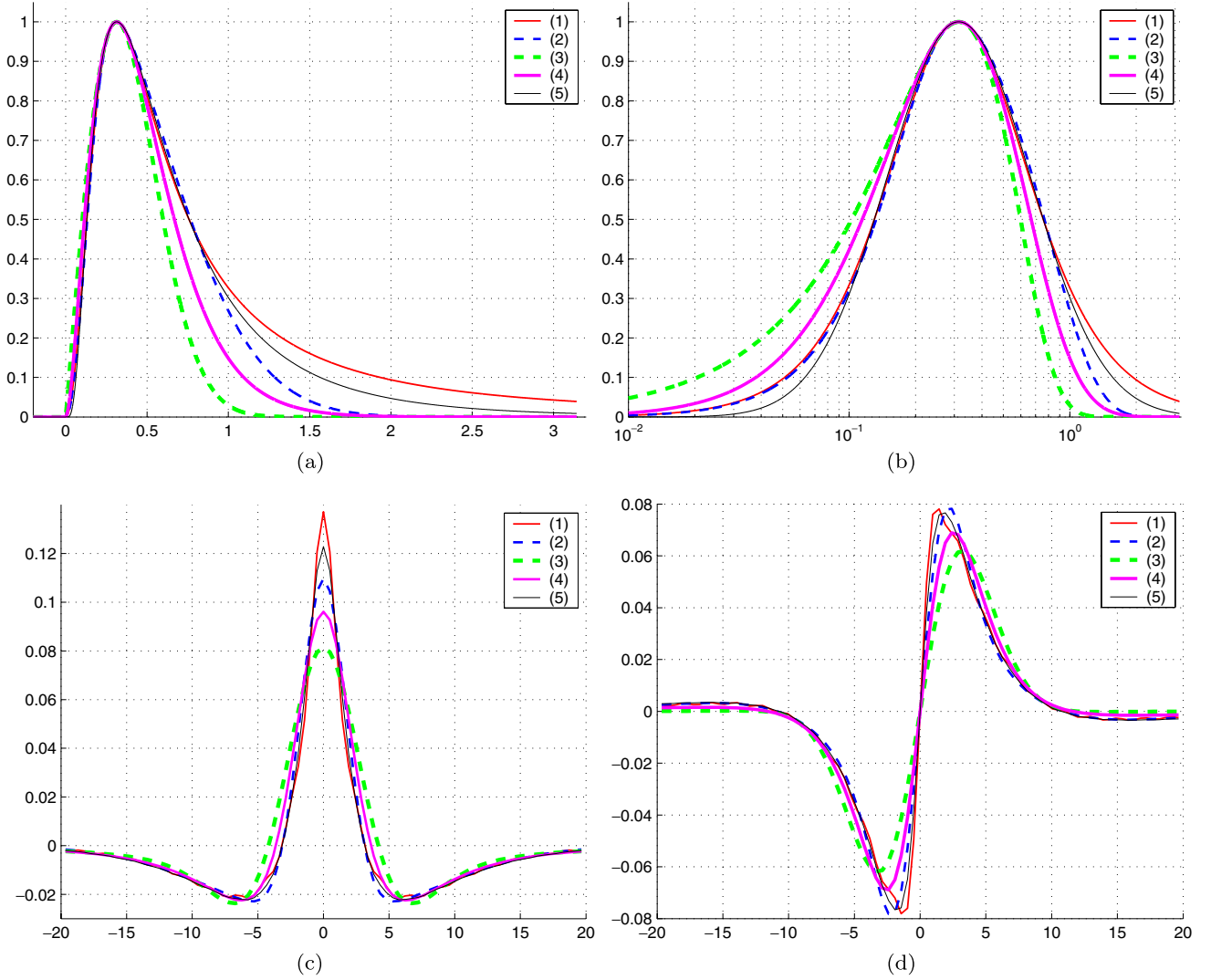


Figure 8. Different quadrature filters in the frequency domain (a) and (b) and in the spatial domain, even part (c) and odd part (d). All the filters have a bandwidth of $\beta = 2.5431$ octaves and a peak tuning frequency for $\lambda_0 = 20$. Deriche filter with $a = 2$ (Red solid curves (1)); DoG filter with $\gamma = 0.215$ (Blue dashed curves (2)); Gaussian derivative with $a = 1.0326$ (Green dashed curves (3)); Cauchy filters with $a = 1.8597$ (Magenta solid curves (4)) and log-Gabor filters with $\kappa_\beta = 0.4730$ (Black solid curves (5)).

for example in feature detection using Phase Congruency and the linearity of phase with respect to a space perturbation is more important in velocity estimation. It is also desirable to know the error in measurement of phase due to the presence of noise. This could for example be taken into account in a regularisation framework of the estimation/detection.

Fleet and Jepson [24] studied the phase behaviour of quadrature filters with respect to a scale-space perturbation (Δx , $\Delta \lambda$) in the neighbourhood of a scale-space point $\mathbf{p}_0(0, \lambda_0)$ in the presence of white Gaussian noise. A short summary of their main results is given in Appendix A. The authors derived an approximation to the

expected mean $\mu(z)$ of phase difference $\Delta \phi$, and an upper bound of the expected mean variation $b(z)$ about the mean phase difference $\mu(z)$. $\mu(z)$ is independent of the bandwidth of the filters but $b(z)$ is not. Phase is perfectly stable when $\Delta \phi$ is constant with respect to changes in scale $\Delta \lambda$; and it is linear with respect to spatial position when $\Delta \phi$ is a linear function of Δx (i.e stability through scale and linearity through space near \mathbf{p}_0). Earlier, in this paper, we provided analytical expressions for phase behaviour for the Gabor and Cauchy kernels, and we studied the others numerically.

Figure 10 shows the level contours of the mean phase difference $\mu(z)$, for five families: Deriche, DoG,

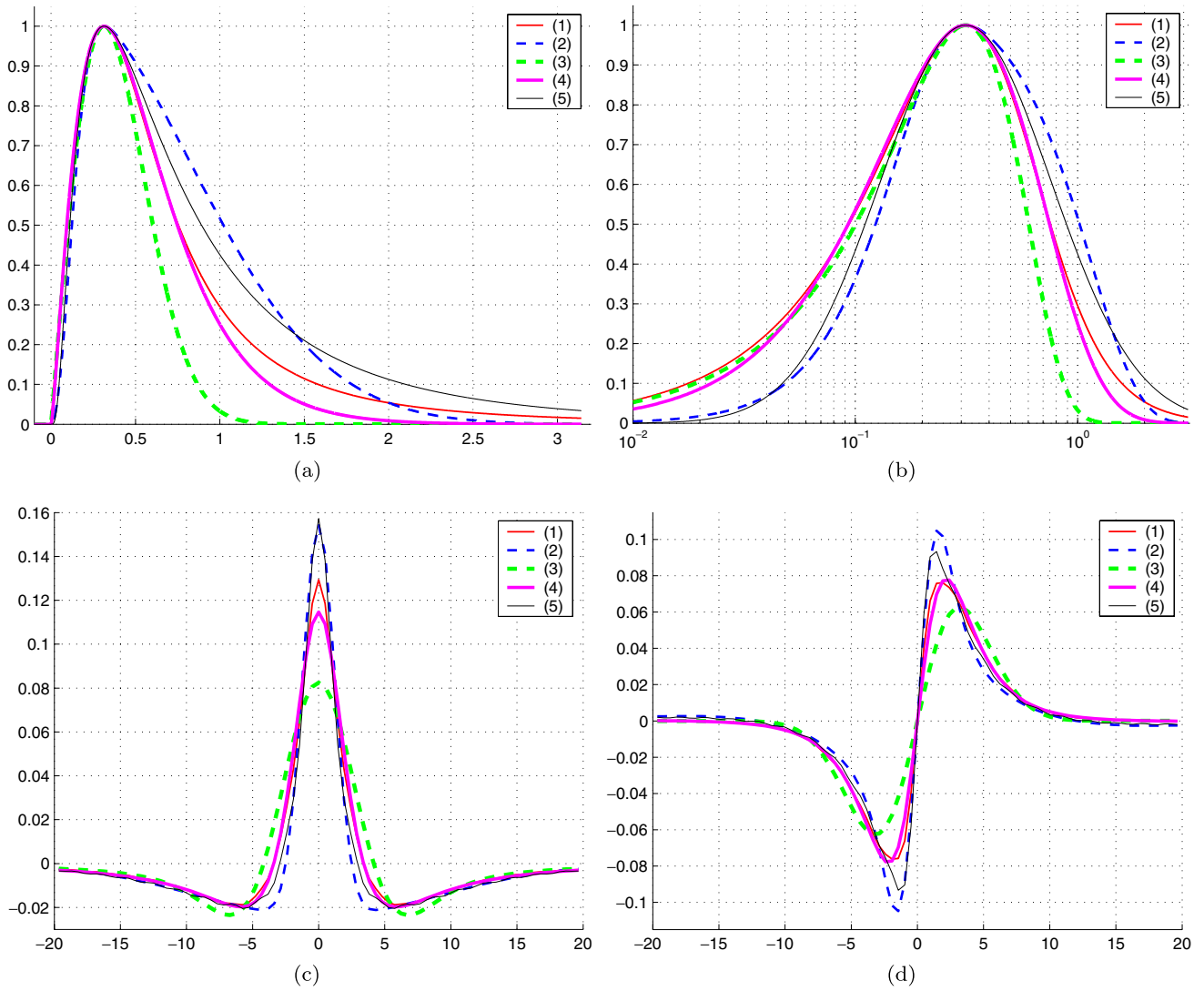


Figure 9. Quadrature filters in the frequency domain (a) and (b) in the spatial domain, even part (c) and odd part (d). All the filters have a bandwidth of $\beta = 3.0096$ octaves (except Gaussian derivative $\beta = 2.5902$) and a peak tuning frequency for $\lambda_0 = 20$. Deriche filter with $a = 1$ (Red solid curves (1)); DoG filter with $\gamma = 0.1350$ (Blue dashed curves (2)); Gaussian derivative with $a = 1$ (Green dashed curves (3)); Cauchy filters with $a = 1.3486$ (Magenta solid curves (4)) and log-Gabor filters with $\kappa_\beta = 0.4124$ (Black solid curves (5)).

Cauchy and log-Gabor for $\beta = 3.0096$ and the 1st derivative of a Gaussian for $\beta = 2.5902$. In theory $\mu(z)$ is independent of the bandwidth of the filters, if the self-similarity condition A.7 is satisfied. This figure illustrates that phase is stable over scale for all the families. However the linearity through space is not perfect for some of them for this bandwidth. This is illustrated in Fig. 10 (Profiles), showing $\mu(z)$ as a function of Δx when $\lambda_1 = \lambda_0$.

We find that phase is perfectly linear only for the 1st derivative of a Gaussian when $-\frac{\lambda_0}{2} \leq x_1 \leq \frac{\lambda_0}{2}$; Cauchy filters and the Deriche filter have a similar behaviour and are close to linear in this range. However the phase is still linear in the vicinity of \mathbf{p}_0 for all filters. The DoG

and log-Gabor kernels have the largest phase distortion. This is a consequence of the tailed nature of the two families for large bandwidths. Inspecting the errors on the mean estimation shown in Fig. 11, we notice that it is for these two families that we have the largest error bound for a spatial perturbation and the smallest for a scale perturbation. This observation, as well, is a consequence of the tailed nature of the log-Gabor and the DoG kernels. As mentioned by Fleet and Jepson [24], as the bandwidth increases, the spatial span of the kernels decreases, and therefore there is a decrease in the spatial extent over which phase is linear. On the other hand, an increase in bandwidth implies that filters tuned to nearby scales will overlap to a great

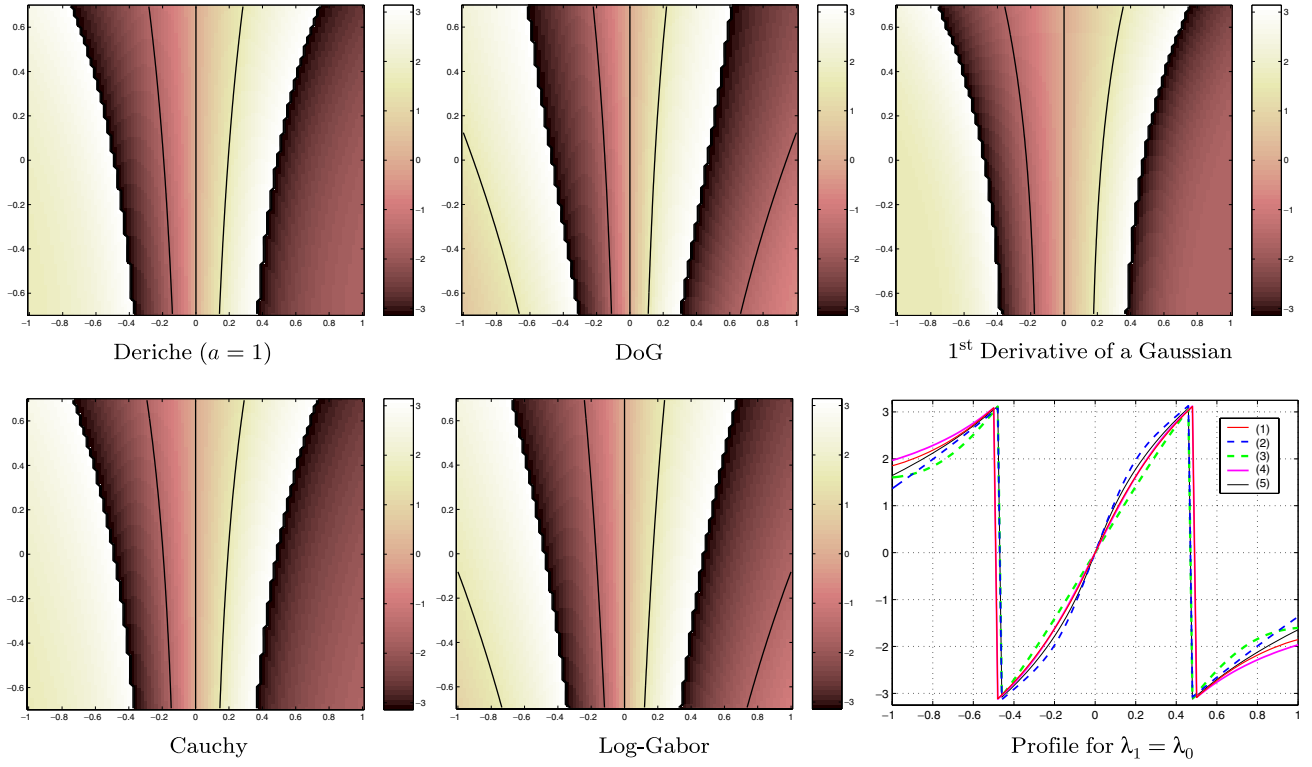


Figure 10. Mean phase difference level contours $\mu(z) = n\pi/2$, with $n = -4, \dots, 4$, near $\mathbf{p}_0 = (0, \lambda_0)$ with a log scale on the vertical axis over two octaves $\frac{\lambda_1}{\lambda_0} \in [\frac{1}{2}, 2]$, and spatial position on the horizontal axis, $\frac{x_1}{\lambda_0} \in [-1, 1]$. All the filters have a bandwidth of 3.0096 octaves except for the 1st derivative of a Gaussian.

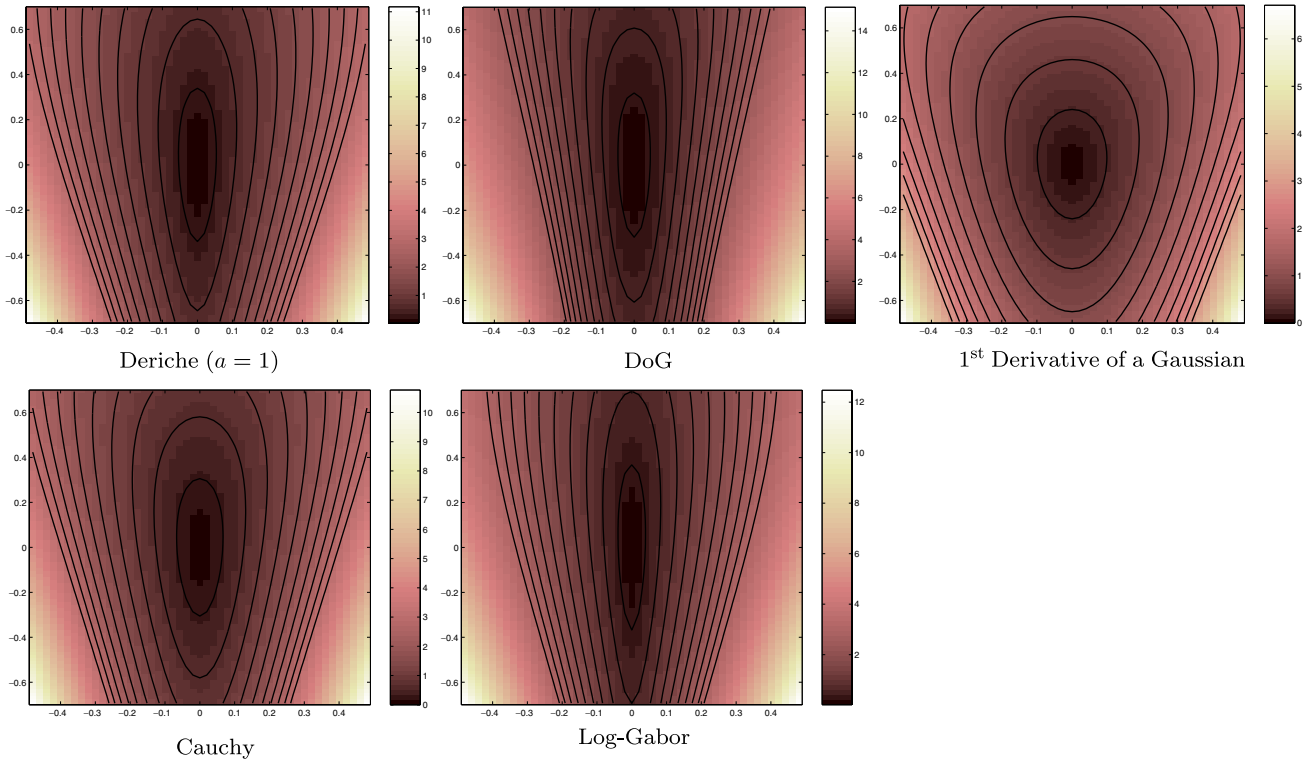


Figure 11. Level contours $b(z) = 0.3n$, with $n = 1, \dots, 10$, corresponding to $\mu(z)$ shown in Fig. 10. All the filters have a bandwidth of 3.0096 octaves except for the 1st derivative of a Gaussian.

extent, and therefore phase is stable for larger scale perturbations. The 1st derivative of a Gaussian has a smaller bandwidth in this experiment, hence, phase is linear for a larger extent. Regarding the other kernels, the larger the extent of the kernel in the Fourier domain, the less is the extent of phase linearity.

Notice that there is a trade-off for the bandwidth between the scale perturbation on the mean phase error $b(z)$, and the spatial one. We will come back to this point when analysing stability for small bandwidths below.

Figure 12, shows the same measures $\mu(z)$ and $b(z)$ for a smaller bandwidth $\beta = 2.5431$ octaves (corresponds to the bandwidth of Ziou's line detector). We notice the same observations regarding the error on the expected mean phase as before. Notice that as the bandwidth of the filters decreases, the expected deviation from $\mu(z)$ is reduced for spatial perturbations and increases for scale perturbations. This is not the case for the Deriche filter, for which the effect is slightly the opposite (see for example the difference between the DoG filter and the Deriche filter). Indeed, as we already mentioned in the last section (see Fig. 8), the Deriche filter has the largest aliasing for this bandwidth (larger than for $\beta = 3.0069$, i.e $a = 1$). This is due to the decay of the Fourier transform of this family ($\omega^{(a-4)}$). Regarding the phase linearity of the filters, as expected, we notice that the extent over which phase is linear has increased.

The phase behaviour for a bandwidth of $\beta = 1.8$ octaves is shown in Fig. 13 for: Gabor kernels with zero DC correction [3], DoG, derivative of a Gaussian, Cauchy and log-Gabor kernels. We did not show the Deriche filters result as the smallest bandwidth of this family is 2.54 octaves. For $\beta = 1.8$, all the filters are perfectly linear for spatial perturbations of more than half of a wavelength. Notice the similar behaviour in terms of the expected deviation $b(z)$ for Gabor, DoG and the Gaussian derivative filters. Notice that a Gaussian derivative filter of a bandwidth of 1.8 octaves corresponds to $a = 1.92$ and a DoG filter of the same bandwidth has a value of 0.7 for γ . Hence, the similarity between the DoG filter and the derivative of a Gaussian filter is expected for this bandwidth, as we know that the 2nd derivative of a Gaussian is well approximated by the DoG filter for $\gamma \approx 0.9$ and that the shape variability of the DoG filter is very low when $\gamma > 0.5$ (see Fig. 4).

Finally, Fig. 14 shows $\mu(z)$ and $b(z)$ for a bandwidth of $\beta = 1$ octave for: Gabor kernels with ([3] and [65]) as zero DC correction, DoG, derivative of a Gaussian,

Cauchy and log-Gabor filters. All the families have a linear phase over the range of x_1 . All the filters have a comparable errors on the estimation of the mean phase. Notice, the circular shape of the level contours of $b(z)$. As we mentioned before, for small bandwidths, the expected deviation from the mean phase for spatial perturbations of the input is reduced and there is an increase of the spatial extent over which phase is linear. On the other hand, there is an increase of the error bound for scale perturbations which decreases the extent of phase stability.

3.2. Uncertainty Principle

The uncertainty principle criteria is a measure of the scale-space localisation of the filters. Gabor functions are known to be optimal in this sense [see e.g. 51, Chap. 2]. Figure 15 shows the standard deviations of the energy of the filters in the spatial domain σ_f and in the frequency domain σ_F and the uncertainty principle, which is the product of both, $\sigma_f \sigma_F$; all as a function of the bandwidth of the filters and for a peak tuning frequency for $\lambda_0 = 20$. σ_f and σ_F are dependent on λ_0 but not the product. The families used in this experiment are: Cauchy filters (1), Gabor with Ronse zero DC correction using $\tilde{\sigma} = \sigma$ (2) and $\tilde{\sigma} = \check{\sigma}$ (3), DoG filters (4), Gaussian derivative filters (5), log-Gabor (6) and Sarkar and Boyer filters for best $\Gamma\Sigma$ (i.e Deriche filter) (7) and best $\Gamma\Sigma\Omega_S$ (8).

Inspecting Fig. 15(a), which shows the spatial standard deviations of the filter's energy as a function of the bandwidth, we notice that all the families have a similar spatial extent for small bandwidths ($\beta < 1.5$). For larger bandwidths, the Gaussian derivative and the corrected Gabor wavelets have the largest σ_f . The three following families: Cauchy, log-Gabor and the DoG filters have a very similar σ_f (the DoG filter being the well localised when $\beta > 2.1$ and the worst for $1.76 < \beta < 1.9$). Interestingly, the dependence of σ_f on the bandwidth for the Deriche filter and the Sarkar and Boyer filter is very low.

In the Fourier domain, the corrected Gabor filters have the best localisation followed by the Gaussian derivative family (see Fig. 15(b)). The behaviour of the Cauchy filters is between the Gaussian derivatives and the log-Gabor, which has the worst frequency localisation. The DoG filter has a better localisation than Cauchy kernels for $1.76 < \beta < 2$ and is worse when $\beta > 2$. Here too, we notice that the dependence of σ_F on the bandwidth for Deriche filters and

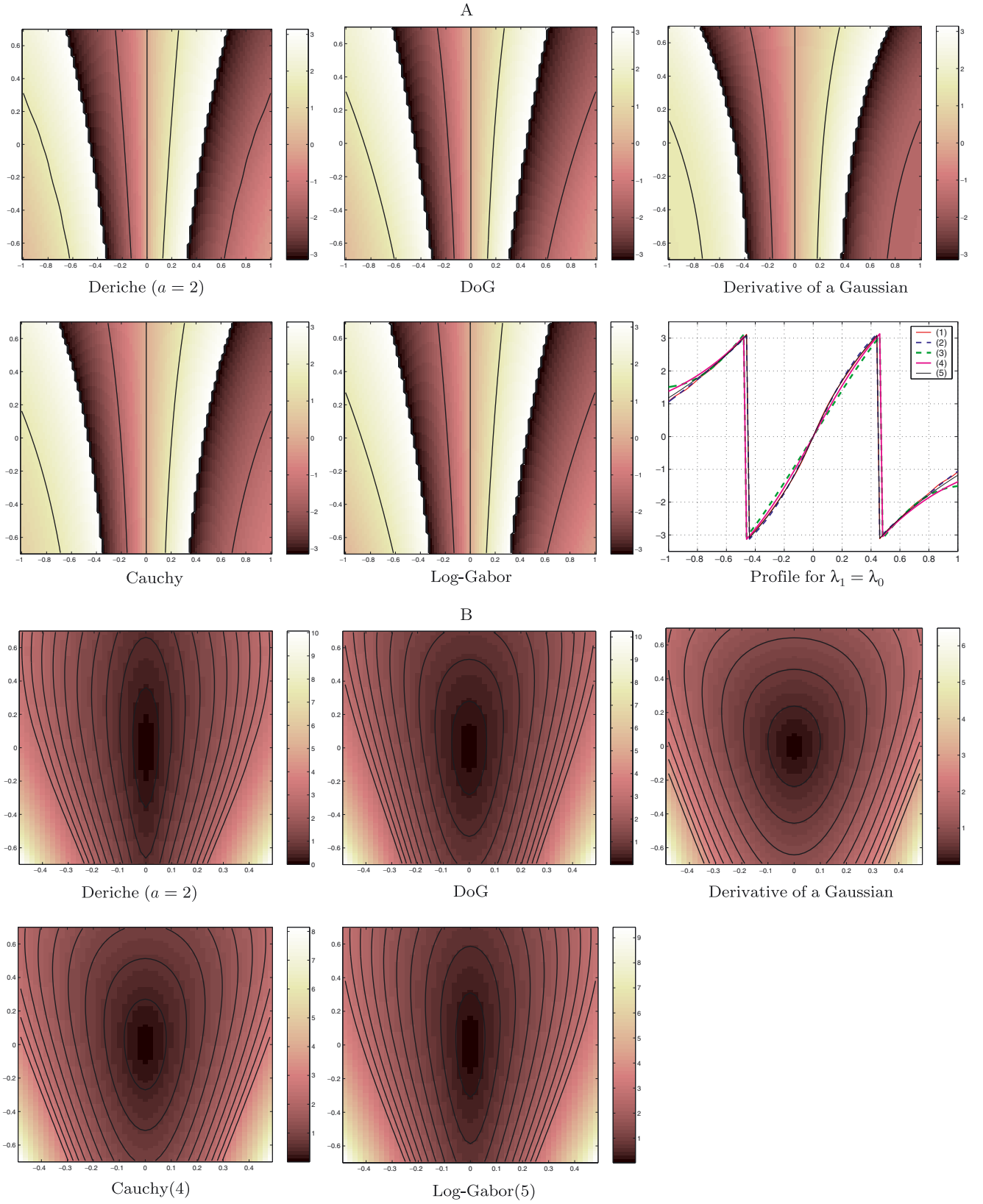


Figure 12. Mean phase difference level contours $\mu(z) = n\pi/2$, with $n = -4, \dots, 4$ (A) and level contours of $b(z) = 0.3n$, with $n = 1, \dots, 10$ (B), near $\mathbf{p}_0 = (0, \lambda_0)$ with a log scale on the vertical axis over two octaves $\frac{\lambda_1}{\lambda_0} \in [\frac{1}{2}, 2]$, and spatial position on the horizontal axis, $\frac{x_1}{\lambda_0} \in [-1, 1]$. All the filters have a bandwidth of 2.5431 octaves.

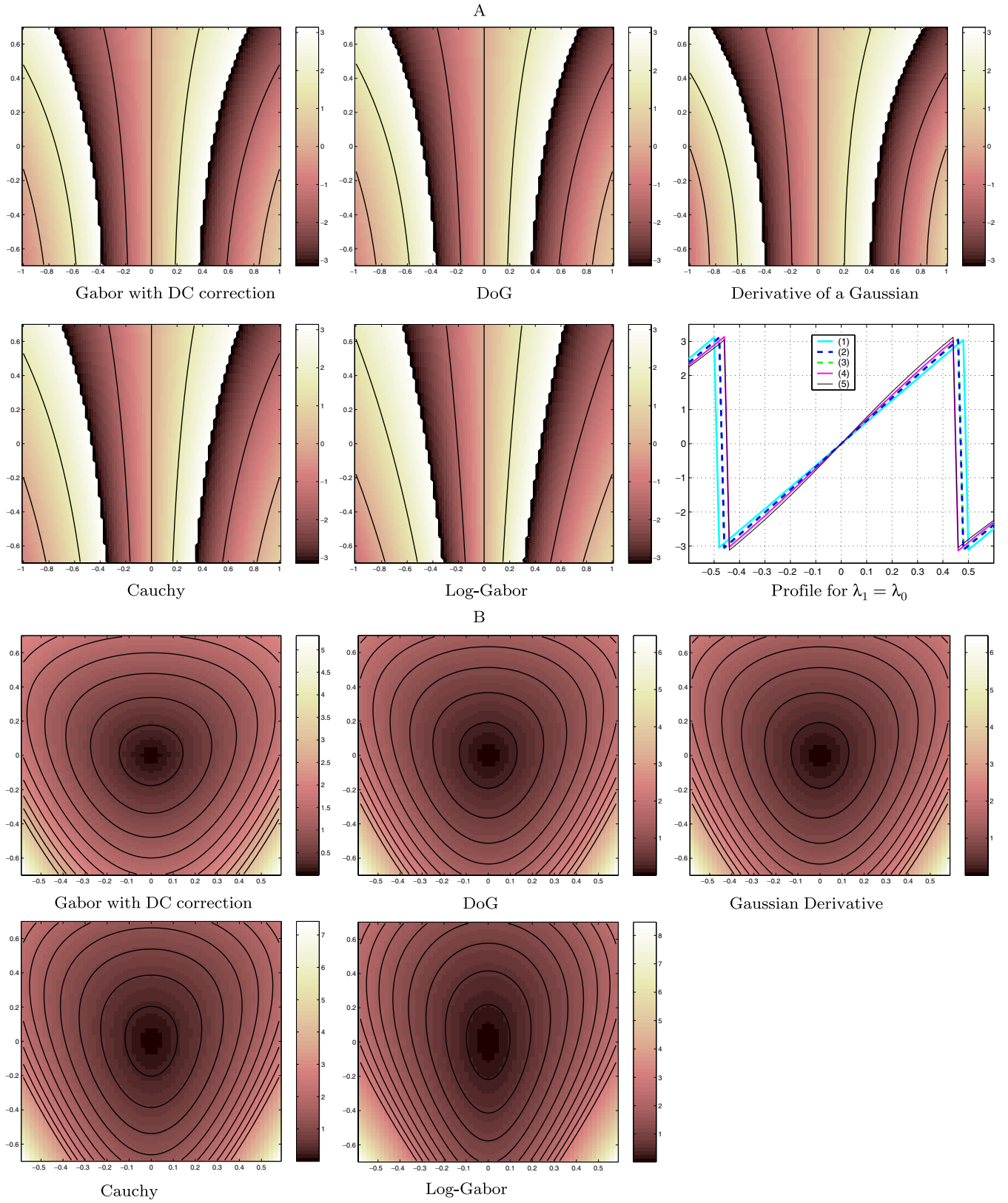


Figure 13. Mean phase difference level contours $\mu(z) = n\pi/2$, with $n = -4, \dots, 4$ (A), and Level contours $b(z) = 0.3n$, with $n = 1, \dots, 10$ (B), near $\mathbf{p}_0 = (0, \lambda_0)$ with a log scale on the vertical axis over two octaves $\frac{\lambda_1}{\lambda_0} \in [\frac{1}{2}, 2]$, and spatial position on the horizontal axis, $\frac{x}{\lambda_0} \in [-1, 1]$. The Gabor zero DC correction is obtained with [3]. All the filters have a bandwidth of $\beta = 1.8$.

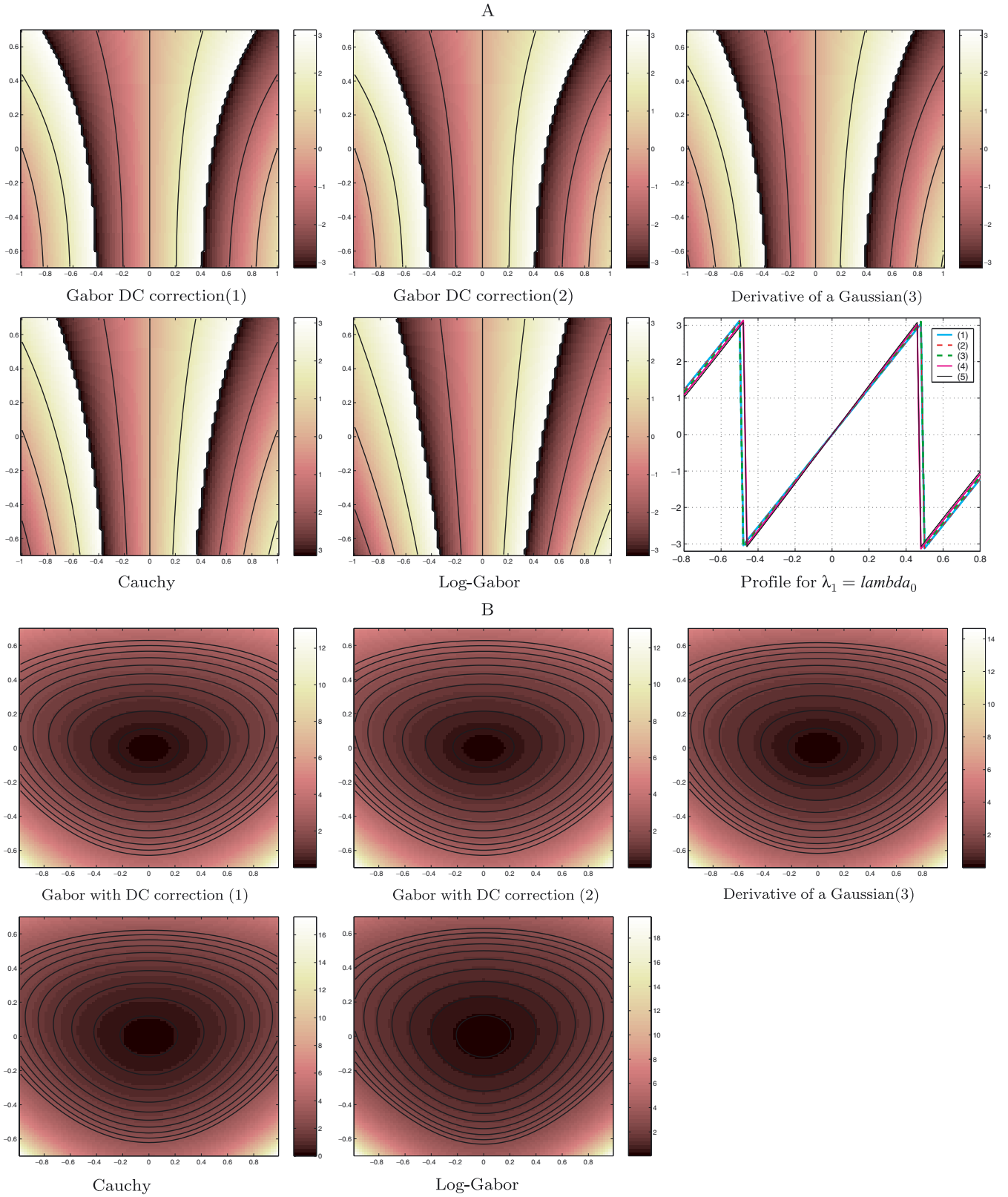


Figure 14. Mean phase difference level contours $\mu(z) = n\pi/2$, with $n = -4, \dots, 4$ (A), and Level contours $b(z) = 0.3n$, with $n = 1, \dots, 10$ (B), near $\mathbf{p}_0 = (0, \lambda_0)$ with a log scale on the vertical axis over two octaves $\frac{\lambda_1}{\lambda_0} \in [\frac{1}{2}, 2]$, and spatial position on the horizontal axis, $\frac{x_1}{\lambda_0} \in [-1, 1]$. Gabor (with [3] as zero DC correction) (1); Gabor (with [65] as zero DC correction) (2). All the filters have a bandwidth of $\beta = 1$.

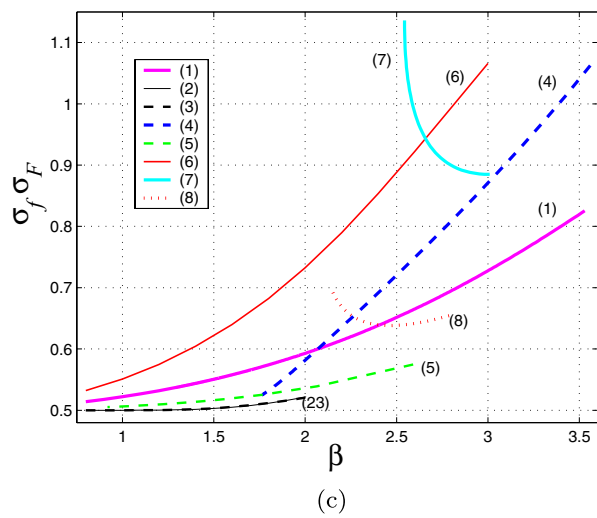
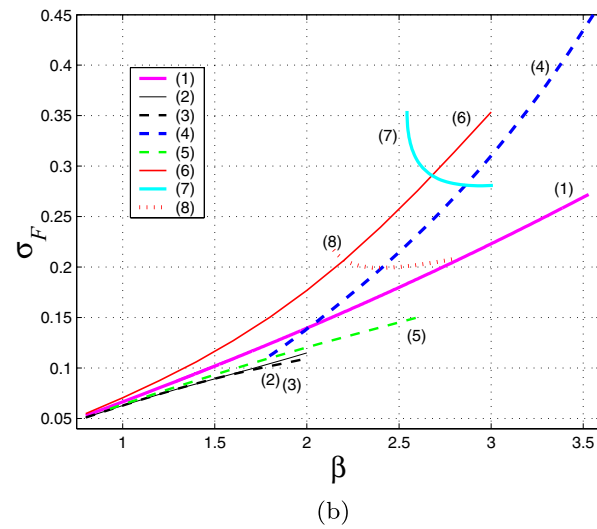
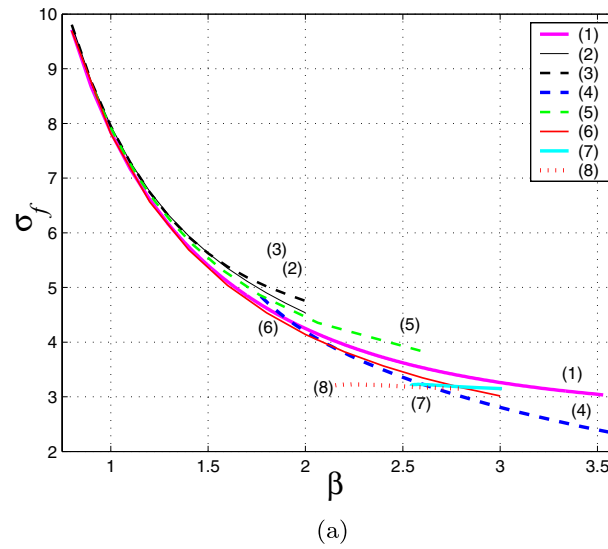


Figure 15. Uncertainty principle for the quadrature pairs for $\lambda_0 = 20$. Cauchy filters (solid Magenta curves (1)), Gabor with Ronse zero DC correction using $\tilde{\sigma} = \sigma$ (solid Black curves (2)) and $\tilde{\sigma} = \tilde{\sigma}$ (dashed Black curves (3)), DoG filters (dashed Blue curves (4)), Gaussian derivative filters (dashed Green curves (5)), log-Gabor (solid Red curves (6)) and Sarkar and Boyer filters for best $\Gamma\Sigma$ (i.e Deriche filter) (solid Cyan curves (7)) and best $\Gamma\Sigma\Omega_S$ (dots Red curves (8)).

Sarkar and Boyer family is opposite to the dependence of other families. This result is expected as, for example, we already pointed out that the Deriche family has more aliasing for smaller bandwidths. This is because the bandwidth is inversely proportional to the order a of the derivative of the smoothing function and the decay of its Fourier transform is of order $\omega^{(a-4)}$. Notice that the Sarkar and Boyer filter has a far better frequency localisation than the Deriche one, especially for smaller bandwidths.

The relative behaviour of the different families regarding the uncertainty principle shown in Fig. 15(c) is similar to their relative behaviour regarding σ_F . The zero DC corrected Gabor filter is close to the limit ($\frac{1}{2}$) almost everywhere, followed by the Gaussian derivatives filter. Because of their tailed nature, the log-Gabor wavelets are the farthest from the limit with the Deriche filter for $2.54 < \beta < 2.66$ (i.e $a \in [1.46, 2]$); the Sarkar and Boyer filters seems to have the best scale-space localisation for $2.5902 < \beta \leq 2.7875$ (basically for a bandwidth greater than the 1st derivative of a Gaussian until the bandwidth of Sarkar and Boyer's edge operator). For very large bandwidths, Cauchy kernels are the closest to the limit.

In the case of feature detection we can locate Canny's, Deriche and Sarkar and Boyer's edge operators as well as Ziou's line detector on these curves (Fig. 15, curves (5), (7), (8) and (7) respectively); these operators are optimal in the sense of Canny's criteria. This gives some insight on the performance of the other families, in the sense of Canny. Recently, the log-Gabor filters have been proposed as the choice of quadrature pair for feature detection. Spatial localisation of the filters is of great interest. On inspection of Fig. 15(a), we notice that the DoG filter has better spatial localisation than log-Gabors for a bandwidth greater than 2.1 octave and the Sarkar and Boyer filters are far better in their range of definition. When $\beta < 2$ log-Gabors are better than the DoG but the Cauchy filter has a comparative localisation. Fig. 15(b) shows that the DoG and the Cauchy family have always a better frequency localisation than log-Gabor kernels and Fig. 15(c) shows that the two families have a better uncertainty principle than the log-Gabor family. All three families are invariant to a grey level shift and to an additive ramp. Finally, note that the Deriche edge operator is closer to the DoG filter, and the Sarkar and Boyer edge operator is closer to the Cauchy family. Hence, for edge detection log-Gabor kernels are not the best choice.

However, for line detection, log-Gabor filters can have good performance as they are the closest to Ziou's optimal line detector. However, we found that the DoG filter for larger bandwidths has similar performance to the Ziou line detector. In addition Ziou's line detector has poor spectral localisation and the largest aliasing.

Our results suggests that the DoG filter is the best to use for combined line and edge detection. The Cauchy family is an interesting alternative, and we have the analytical expressions of the filters in the spatial and the frequency domain. Further experiments are needed to investigate their use in more detail.

4. Concluding Remarks

We began by stressing the importance of phase information in computer vision, then provided the reader with a short summary of the definition of the local phase in one dimension, and its extension to multi-dimensional signals. In practice, the estimation of the local phase and the related quantity, local energy, uses a pair of band-pass quadrature filters. Several families of quadrature pairs have been proposed and applied in the literature, in a number of different applications. Most authors have not provided a reasonable justification for the use of a particular family apart from simplicity of use or the satisfaction of the zero DC condition. This is particularly the case in feature detection where, for example, several authors have not considered the invariance of the filters to an additive ramp (smooth gradient). This property is desirable and has been noted in classical feature detection (see e.g. [7]). Our analysis attempts to put the justification on a sound foundation.

Specifically, in the paper, we have studied mathematically the most common quadrature filters in the literature. We have provided either a mathematical expression or a numerical estimation for most filter properties. Key results and properties are summarised in Table 3.

In the case of feature detection, we presented the first attempt to use classical edge/line operators for the design of a quadrature pair. Although the main advantage of phase-based feature detection is that it is a unified detection approach (not only for edges or lines), we can learn a lot from the separate behaviour of the even part (symmetric detector, lines for example) and the odd part (asymmetric detector, edges for example). Interestingly, we found that most of the classical optimal operators, have large bandwidth and several

authors point out that broad bandwidth filters have a better SNR in phase-based detection.

Our analysis shows that Gabor filters are not a good choice for feature detection. The odd part has poor performance in the sense of Canny in comparison to the 1st derivative of a Gaussian, to Sarkar and Boyer filter and to the Deriche filter; and the even part does not satisfy the zero DC correction and hence needs correction. The zero DC correction is possible but the filters are still not truly in quadrature, except if the negative frequencies are zeroed in the case of a Fourier domain implementation. Gabor filters are sensitive to smooth gradients.

We also concluded that log-Gabor kernels are not a very good choice for feature detection. As we showed, the DoG family and the Cauchy family have better properties. The Sarkar and Boyer's filter has very good properties, for example for a bandwidth close to 2 octaves, this family has very good spatial localisation. In the case that we are interested only in one type of features (edges or lines), a natural choice is to use the Sarkar and Boyer family which is optimal in the sense of Canny.

Regarding the use of quadrature filters for velocity estimation, we showed that all the families have similar behaviours for small bandwidths (in terms of phase linearity and phase stability). The zero DC corrected Gabor filters could be used. If we want to be invariant to a smooth gradient (in the case where the conservation criteria is violated), the Gaussian derivative family would be a natural choice as they have close behaviour to the Gabor filters in terms of phase stability and the uncertainty principle. The disadvantage of using the Gaussian derivative filter is the loss of the spatial representation of the filters. In this case, the Cauchy family offers an alternative and has the advantage of being invariant to smooth gradients and having analytical expression in the spatial and the Fourier domain.

Appendix A: Phase Stability Given White Gaussian Noise

The stability of phase can be measured in terms of the differences in phase response between two arbitrary neighbouring scale-space locations, $\mathbf{p}_0 = (x_0, \lambda_0)$, $\mathbf{p}_1 = (x_1, \lambda_1)$. We denote the filter response to the input $f(x)$ at scale-space position \mathbf{p}_j by:

$$g_j = h(x - x_j, \lambda_j) \otimes f(x) = \langle h(x_j - x, \lambda_j), f(x) \rangle, \quad (\text{A.1})$$

where $\langle f(x), g(x) \rangle = \int_{-\infty}^{\infty} f(x)^* g(x) dx$ is the inner product.

The phase differences, as a function of the relative scale-space position, $\mathbf{p}_1 - \mathbf{p}_0 = (\Delta x, \Delta \lambda)$ can be written as:

$$\Delta \phi = \arg[g_1] - \arg[g_0]. \quad (\text{A.2})$$

Phase is perfectly stable when $\Delta \phi$ is constant with respect to changes in scale $\Delta \lambda$; and it is linear with respect to spatial position when $\Delta \phi$ is linear function of Δx .

To characterise the stability of phase behaviour through scale-space, Fleet and Jepson [24] examined the response of the filter kernel, $h(x, \lambda)$ and its phase behaviour to stationary, zero mean white Gaussian noise. The authors derived an approximation to the mean phase difference $E[\Delta \phi]$, and an upper bound of the mean variation about the mean $E[|\Delta \phi - E[\Delta \phi]|]$. Given the above noise assumptions:

$$E[\Delta \phi] \approx \mu(z) = \arg[z] \quad (\text{A.3})$$

$$E[|\Delta \phi - E[\Delta \phi]|] \leq b(z) = \frac{\sqrt{1 - |z|^2}}{|z|}, \quad (\text{A.4})$$

where the complex scalar z is a function of scale-space position and is given by:

$$\begin{aligned} z(\mathbf{p}_1 - \mathbf{p}_0) &= z(\Delta x, \Delta \lambda) \\ &= \langle h(x_0 - x, \lambda_0), h(x_1 - x, \lambda_1) \rangle, \end{aligned} \quad (\text{A.5})$$

where the kernel h has to satisfy the two following conditions:

(1) Normalisation condition:

$$\|h(x, \lambda)\| = \langle h(x, \lambda), h(x, \lambda) \rangle = 1. \quad (\text{A.6})$$

(2) Translational invariance and self-similarity across scales (i.e their octave bandwidth remains constant, independent of the scale to which they are tuned):

$$h(x, \lambda) = \frac{1}{\sqrt{\lambda}} h(x/\lambda, 1). \quad (\text{A.7})$$

Appendix B: Optimal Linear Edge Detector Criteria

Consider the problem of detecting an ideal feature (edge or line) $f_F(x)$, centred at $x=0$ using an

observation corrupted by an additive white Gaussian noise with standard deviation $\sigma = 1$. Canny [7, 8] was the first to formulate the detection problem for a linear edge detector $f_O(x)$ as an optimisation problem. As Canny pointed out, the key to the design of an effective edge operator is accurate evaluation of its performance. Canny defined three different criteria to be maximised:

Good Detection: Given the white Gaussian noise assumption, the probability of false detection decreases with increasing the signal-to-noise ratio $\Sigma(f)$ given by:

$$\Sigma(f_O) = \frac{|f_F(x) \otimes f_O(x)|}{\sqrt{\int_{-\infty}^{\infty} f_O^2(x) dx}} = \frac{|\int_{-\infty}^{\infty} f_F(-x) f_O(x) dx|}{\sqrt{\int_{-\infty}^{\infty} f_O^2(x) dx}}. \quad (\text{B.1})$$

Good Localisation: This second criterion is defined as the variance of the maximum of the filter response about the correct location.⁹ Canny used an approximation based on a Taylor's expansion of the filter output:¹⁰

$$\text{for edges} \quad \Lambda(f_O) = \frac{|\int_{-\infty}^{\infty} f'_F(-x) f'_O(x) dx|}{\sqrt{\int_{-\infty}^{\infty} f_O'^2 dx}}. \quad (\text{B.2})$$

$$\text{for lines} \quad \Lambda(f_O) = \frac{|\int_{-\infty}^{\infty} f_F(-x) f''_O(x) dx|}{\sqrt{\int_{-\infty}^{\infty} f_O'^2 dx}}. \quad (\text{B.3})$$

Low Number of Multiple Responses: This criterion was designed to minimise the number of spurious maxima within the operator spatial spread [8]:

$$\Omega_C(f_O) = 2\pi \sqrt{\frac{\int_{-W}^W f_O' dx}{\int_{-W}^W f_O''^2 dx}} = kW, \quad (\text{B.4})$$

where W is the operator width and k is a positive constant.

Sarkar and Boyer [67] point out that the above multiple response criteria must be modified to be applied to IIR filters. In other words we should provide an estimate of the operator width Sarkar and Boyer [67] defined the width of the filter as the normalised root mean square deviation of the function $f_O^2(x)$ from the

origin.

$$\bar{W} = \sqrt{\frac{\int_{-\infty}^{\infty} x^2 f_O^2(x) dx}{\int_{-\infty}^{\infty} f_O^2(x) dx}}. \quad (\text{B.5})$$

Hence, they proposed a new expression for $\Omega(f)$:

$$\Omega_S(f_O) = 2\pi \sqrt{\frac{\int_{-\infty}^{\infty} f_O'^2 dx}{\int_{-\infty}^{\infty} f_O''^2 dx}} \sqrt{\frac{\int_{-\infty}^{\infty} f_O^2(x) dx}{\int_{-\infty}^{\infty} x^2 f_O^2(x) dx}}, \quad (\text{B.6})$$

Then, the optimality of the detector is defined with respect to the maximisation of the following quantity [8]:

$$C_c(f_O) = \Gamma(f_O) \Sigma(f_O) \Omega(f_O). \quad (\text{B.7})$$

For an ideal step edge (B.7) is given by:

$$C_c(f_O) = \frac{|\int_{-\infty}^0 f_O(x) dx| |f'_O(0)|}{\sqrt{\int_{-\infty}^{\infty} f_O^2(x) dx} \sqrt{\int_{-\infty}^{\infty} f_O''^2 dx}}. \quad (\text{B.8})$$

Acknowledgments

This work was funded by the EC project ADEQUATE (IST: 1999-10837) and the UK EPSRC funded IRC in Medical Image Analysis and Signals (MIAS).

Notes

1. Throughout, we mean “local” or “instantaneous” . . . When 1-D signals are discussed, the “space” variable may be called “time”.
2. Throughout, we use the following convention: if $f(x)$ is a 1D function then $F(\omega)$ is its Fourier transform.
3. Normally, energy is define as the amplitude-squared [26, 65], however, some authors (see for e.g. [40, 57, 64]), define the energy as being the amplitude.
4. Of course, this is not an exhaustive list of quadrature filters in the literature. We did not include, for example, *shiftable* [70] filters as this property implies their use in a multiscale framework as well as the robust quadrature filters proposed in [54].
5. Note: the use of these two normalisation conditions is a recurrent theme throughout.
6. Strictly speaking, log-Gabor filters are invariant to an additive polynomial function of order $n < \infty$.
7. Mathematically, the 2nd derivative of a Gaussian is well approximated by the DoG filter when $\sigma' \rightarrow \sigma$. However, the DoG filter preferred by Marr and Hildreth [53] is for $\sigma = 1.6 \sigma'$.

8. Of course the optimality is debatable as for example we found that the even DoG filter (when $\gamma \leq 0.05$) has better performance than Ziou's line detector, in the sense of Canny, using the following line model: $F(x) = \text{rect}((x - 0.25)/0.5)$.
9. Although the agreement [5, 67] and the disagreement [73], this is the most popular localisation measure in this context.
10. The Taylor expansion of the filters output are different for an edge detector and a line detector because the edge detector is odd and the line detector is even. See Canny [8] for the details.

References

1. T. Aach, A. Kaup, and R. Mester, "On texture analysis: Local energy transforms versus quadrature filters," *Signal Processing*, Vol. 45, pp. 173–181, 1995.
2. J.L. Barron, D.J. Fleet, and S.S. Beauchemin, "Performance of optical flow techniques," *Int. J. Computer Vision*, Vol. 12, pp. 43–77, 1994.
3. D. Boukerroui, J.A. Noble, and M. Brady, "On Gabor zero DC correction," Technical Report, MVL, Oxford University, 2002. <http://www.robots.ox.ac.uk/~djamal/>.
4. A.C. Bovik, M. Clark, and W.S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. PAMI*, Vol. 12, No. 1, pp. 55–73, 1990.
5. K.L. Boyer and S. Sarkar, "On the localisation performance measures and optimal edge detection," *IEEE Trans. PAMI*, Vol. 16, No. 1, pp. 106–108, 1994.
6. T. Bülow, "Hypercomplex spectral signal representations for the processing and analysis of images," Ph.D. dissertation, Christian Albrechts, Univ., Keil, Germany, 1999.
7. J.F. Canny, "Finding edges and lines in images," MIT Artif. Intell. Lab., Massachusetts Inst. Technol., Cambridge, MA, A.I. Memo 720, 1983.
8. J.F. Canny, "A computational approach to edge detection," *IEEE Trans. PAMI*, Vol. 8, pp. 679–714, 1986.
9. D. Casasent and J.S. Smokelin, "Real, imaginary, and clutter Gabor filter fusion for detection with reduced false alarms," *Optical Engineering*, vol. 33, pp. 2255–2263, 1994.
10. C.-M. Chen, H.H.-S. Lu, and K.-C. Han, "A textural approach based on Gabor functions for texture edge detection in ultrasound images," *Ultrasound in Med. & Biol.*, Vol. 27, No. 4, pp. 515–534, 2001.
11. R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.E. Jeffrey, and D.E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, Vol. 5, pp. 329–359, 1996. <http://pineapple.apmaths.uwo.ca/~rmc/papers/LambertW/index.html>.
12. J.G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Research*, Vol. 10, pp. 847–856, 1980.
13. R. Deriche, "Optimal edge detection using recursive filtering," in *Proc. Int. Conf. Computer Vision*, London, UK, 1987, pp. 167–187.
14. R. Deriche, "Fast algorithms for low-level vision," *IEEE Trans. PAMI*, Vol. 12, No. 1, pp. 78–86, 1990.
15. D. Dunn and W.E. Higgins, "Optimal Gabor filters for texture segmentation," *IEEE Trans. Image Processing*, Vol. 4, No. 7, pp. 947–964, 1995.
16. M. Elad, P. Teo, and Y. Hel-Or, "Optimal filters for gradient-based motion estimation," in *Proc. Int. Conf. Computer Vision*, Corfu, Greece, 1999, pp. 559–565.
17. I.H. Elder and S.W. Zucker, "Local scale control for edge detection and blur estimation," *IEEE Trans. PAMI*, Vol. 20, No. 7, pp. 699–716, 1998.
18. M. Felsberg and G. Sommer, "A new extension of linear signal processing for estimating local properties and detecting features," *DAGM Symposium Mustererkennung*, Kiel, Germany, 2000, pp. 195–202.
19. M. Felsberg and G. Sommer, "The multidimensional isotropic generalisation of quadrature filters in geometric algebra," in *Proc. Int. Workshop on Algebraic Frames for the Perception-Action Cycle*, Kiel, Germany, 2000, pp. 175–185.
20. M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Processing*, Vol. 49, No. 12, pp. 3136–3144, 2001.
21. M. Felsberg and G. Sommer, "Scale adaptive filtering derived from the Laplace equation," in *Lecture Notes in Computer Science, 23rd DAGM Symp. Mustererkennung*, 2191, B. Radig and S. Florczyk (Eds.), Heidelberg, Germany, 2001.
22. D.J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Am.*, Vol. A, No. 4, pp. 2379–2394, 1987.
23. D.J. Fleet and A.D. Jepson, "Computational of component image velocity from local phase information," *Int. J. Computer Vision*, Vol. 5, pp. 77–104, 1990.
24. D.J. Fleet and A.D. Jepson, "Stability of phase information," *IEEE Trans. PAMI*, Vol. 15, No. 12, pp. 1253–1268, 1993.
25. W.T. Freeman and D. Adam, "The design and the use of steerable filters," *IEEE Trans. PAMI*, Vol. 13, No. 9, pp. 891–906, 1991.
26. D. Gabor, "Theory of communication," *J. IEE*, Vol. 93, pp. 429–457, 1946.
27. G.H. Granlund and H. Knutsson, *Signal Processing for Computer Vision*, Kluwer Academic Publishers: Dordrecht, 1995.
28. S.L. Hahn, *Hilbert Transform in Signal Processing*, Artech House: Boston, MA 1996.
29. E.R. Hancock, "Resolving edge-line ambiguities using probabilistic relaxation," in *Proc. IEEE Computer Vision and Pattern Recognition Conference*, 1993, pp. 300–306.
30. F. Heitger, L. Rosenthaler, R. von der Heydt, E. Peterhans, and O. Kübler, "Simulation of neural contour mechanisms: From simple to the end-stopped cells," *Vision Research*, Vol. 32, pp. 963–981, 1992.
31. F. Heitger, "Feature detection using suppression and enhancement," Tech. Rep., Image Science Lab, ETH-Zurich, Switzerland, NR. 163, 1995.
32. R. Highnam and M. Brady, *Mammographic Image Analysis*, Kluwer, 1999.
33. C.-L. Huang and Y.-T. Chen, "Motion estimation method using a 3D steerable filter," *Image and Vision Computing*, Vol. 13, No. 1, pp. 21–32, 1995.
34. M. Idrissa and M. Achery, "Texture classification using Gabor filters," *Pattern Recognition Letters*, Vol. 23, pp. 1095–1102, 2002.
35. A.K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, Vol. 23, No. 12, pp. 1167–1186, 1991.
36. A.D. Jepson and M. Jenkin, "Fast computation of disparity from phase differences," in *Proc. IEEE Computer Vision and Pattern Recognition Conference*, San Diego, 1989, pp. 398–403.

37. L.M. Kennedy and M. Basu, "A Gaussian derivative operator for authentic edge detection and accurate edge localization," *Int. J. Pattern Recognition and Artificial Intelligence*, Vol. 13, pp. 367–380, 1999.
38. S.A. Klein and D.M. Levi, "Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation," *J. Opt. Soc. Am., Ser. A*, Vol. 2, No. 7, pp. 1170–1190, 1985.
39. A.K. Klein, F. Lee, and A.A. Amini, "Quantitative coronary angiography with deformable spline models," *IEEE Trans. Medical Imaging*, Vol. 16, No. 5, pp. 468–482, 1997.
40. P. Kovess, "Image feature from phase congruency," *Videre: Journal of Computer Vision Research*, Vol. 1, No. 3, pp. 1–26, 1999.
41. P. Kruizinga and N. Petkov, "Nonlinear operator for oriented texture," *IEEE Trans. Image Processing*, Vol. 8, No. 10, pp. 1395–1407, 1999.
42. P. Kruizinga and N. Petkov, "Nonlinear operator for blob texture segmentation," in *Proc. NSIP'99, IEEE Workshop on Nonlinear Signal Processing*, II, A.S. Cetin, et al. (Eds.), Antalya, Turkey, 1999, pp. 881–885.
43. P. Kruizinga, N. Petkov, and S.E. Grigorescu, "Comparison of texture features based on Gabor filters," in *Proc. of the 10th International Conference on Image Analysis and Processing*, Venice, Italy, 1999, pp. 142–147.
44. C. Kuglin and D. Hine, "The phase correlation image alignment method," *IEEE Int. Conf. Cybern. Society*, 1975, pp. 163–165.
45. K.G. Larkin, D.J. Bone, and M.A. Oldfield, "Natural demodulation of two-dimensional fringe patterns: I. General background of the spiral phase quadrature transform," *J. Opt. Soc. Am. A*, Vol. 18, No. 8, pp. 1862–1870, 2001.
46. T.S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Trans. PAMI*, Vol. 18, No. 10, pp. 959–971, 1996.
47. J.A.F. Leite and E.R. Hancock, "Statistically combining and refining multichannel information," in *Progress in Image Analysis and Processing III*, S. Impedovo (Ed.), World Scientific, 1994, pp. 193–200.
48. J. Little, "Accurate early detection of discontinuities," *Vision Interface*, pp. 97–102, 1992.
49. H. Liu, T.-H. Hong, M. Herman, and R. Chellappa, "A general motion model and spatio-temporal filters for computing optical flow," *Int. J. Computer Vision*, Vol. 22, No. 2, pp. 141–172, 1997.
50. D.G. Lowe, "Object recognition from local scale-invariant features," in *Proc. Int. Conf. Computer Vision*, 2, Corfu, Greece, 1999, pp. 1150–1157.
51. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
52. S. Marčelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Am.*, Vol. 70, pp. 1297–1300, 1980.
53. D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Royal Society London*, Vol. B207, pp. 187–217, 1980.
54. J.L. Marroquin, J.E. Figueroa, and M. Servin, "Robust quadrature filters," *J. Opt. Soc. Am. A*, Vol. 14, No. 4, pp. 779–791, 1997.
55. K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Proc. Int. Conf. Computer Vision*, 1, Vancouver, Canada, 2001, pp. 525–531.
56. M.C. Morrone and R.A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, Vol. 6, pp. 301–313, 1987.
57. M.C. Morrone, A. Navangione, and D. Burr, "An adaptive approach to scale selection for line and edge detection," *Pattern Recognition Letters*, Vol. 16, pp. 667–677, 1995.
58. M. Mulet-Parada and J.A. Noble, "2D + T acoustic boundary detection in echocardiography," *Medical Image Analysis*, Vol. 4, pp. 21–30, 2000.
59. K. Okajima, "The Gabor function extracts the maximum information from input local signals," *Neural Networks*, Vol. 11, pp. 435–439, 1998.
60. K. Okajima, "Two-dimensional Gabor-type receptive field as derived by mutual information maximization," *Neural Networks*, Vol. 11, pp. 441–447, 1998.
61. D. Reisfeld, "The constrained phase congruency feature detector: Simultaneous localisation, classification and scale determination," *Pattern Recognition Letters*, Vol. 17, pp. 1161–1169, 1996.
62. N. Petkov and P. Kruizinga, "Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells," *Biological Cybernetics*, Vol. 76, No. 2, pp. 83–96, 1997.
63. T. Poggio, H. Voorhees, and A. Yuille, "A regularised solution to edge detection," MIT Artif. Intell. Lab., Massachusetts Inst. Technol., Cambridge, MA, A.I. Memo 833, 1985.
64. B. Robbins and R. Owens, "2D feature detection via local energy," *Image and Vision Computing*, Vol. 15, pp. 353–368, 1997.
65. C. Ronse, "On idempotence and related requirements in edge detection," *IEEE Trans. PAMI*, Vol. 15, No. 5, pp. 484–491, 1993.
66. C. Ronse, "The phase congruence model for edge detection in two-dimensional pictures: A mathematical study," Tech. Rep., Rap95-11, Université de Strasbourg, LSIIT-URA 1871, 1995.
67. S. Sarkar and K.L. Boyer, "On optimal infinite impulse response edge detection filters," *IEEE Trans. PAMI*, Vol. 13, No. 11, pp. 1154–1171, 1991.
68. J. Shen and S. Castan, "An optimal linear operator for edge detection," in *Proc. IEEE Computer Vision and Pattern Recognition Conference*, June, Miami, FL, 1986, pp. 109–114.
69. J. Shen and S. Castan, "An optimal linear operator for step edge detection," *CVGIP: Graphical Models and Image Processing*, Vol. 54, No. 2, pp. 112–133, 1992.
70. E.P. Simoncelli, W.T. Freeman, E.H. Adelson, and D.J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Information Theory, Special Issue on Wavelets*, Vol. 38, No. 2, pp. 587–607, 1992.
71. E.P. Simoncelli, "Design of multi-dimensional derivative filters," in *Proc. IEEE Int. Conf. Image Processing*, Austin, Texas, 1994, pp. 790–794.
72. E.P. Simoncelli and H. Farid, "Steerable wedge filters for local orientation analysis," *IEEE Trans. Image Processing*, Vol. 5, No. 9, pp. 1377–1382, 1996.
73. H.D. Tagare and J.P. deFigueiredo, Reply to "On the localisation performance measure and optimal edge detection," *IEEE Trans. PAMI*, Vol. 16, No. 1, pp. 108–110, 1994.
74. S. Venkatesh and R. Owens, "On the classification of image features," *Pattern Recognition Letters*, Vol. 11, pp. 339–349, 1990.
75. T.P. Weldon, W.E. Higgins, and D.F. Dunn, "Gabor filter design for multiple texture segmentation," *Optical Engineering*, Vol. 35, No. 10, pp. 2852–2863, 1996.
76. T.P. Weldon and W.E. Higgins, "An algorithm for designing multiple Gabor filters for segmenting multi-textured images," in *Proc. IEEE Int. Conf. Image Processing*, Chicago, IL, 1998.

77. L. Wiskott, "Segmentation from motion: combining Gabor- and Mallat-wavelets to overcome the aperture and correspondence problems," *Pattern Recognition*, Vol. 32, pp. 1751–1766, 1999.
78. D. Ziou, "Line detection using an optimal IIR filter," *Pattern Recognition*, Vol. 24, No. 6, pp. 465–478, 1991.
79. D. Ziou and S. Tabbone, "Edge detection techniques: An overview," *Pattern Recognition and Information Analysis*, Vol. 8, pp. 537–554, 1998.
80. D. Ziou, "Finding lines in grey-level images," Tech. Rep., N. 240. Université de Sherbrooke, 1999.



Djamel Boukerroui was borne in Bejaïa (Algeria) in 1972. He received the B.S. degree in 1990, and the M.S. degree in electronics in July 1995 from ENP of Algiers (Algeria). He received the Ph.D degree in Image Processing at CREATIS Laboratory of INSA of Lyon (France) in 2000. From March 2000 to August 2002 he was a Research Assistant at the Medical Vision Laboratory, Department of Engineering Science, University of Oxford where he worked on the analysis of echocardiographic image sequences. Since September 2002, he joined the Department of Information Processing Engineering of Compiègne University of Technology as Maître de Conférences, and he is a member of HEUDIASYC [research unit associated with CNRS (#6599)]. His future research will include, Features extraction, optical flow estimation and Content Based Medical Image Retrieval.



J. Alison Noble D. Phil (Oxon), MA (Oxon), FIEE Professor Alison Noble has 16 years of research experience in image analysis, including 5 years in industry, and has been a faculty member at the University of Oxford for 8 years. Her current main research interests are in low level vision, and specifically its application in ultrasound image analysis and in-vivo soft tissue analysis (of the heart and breast). She has 27 journal publications in computer vision and its application in manufacturing and medicine, around 75 other publications presented at conferences, 3 patents (2 US, 1 UK) and three further UK patents pending. She is on the Editorial Advisory Boards of Medical Image Analysis (Elsevier) and Ultrasound in Medicine and Biology (Elsevier) journals, is a regular reviewer for a number

of other journals and international conferences in her research field, and is Co-Chair of the 18th International Conference on Information Processing in Medical Imaging to be held in Ambleside, UK, in 2003.



Michael Brady FRS, FREng BP Professor of Information Engineering at the University of Oxford. Professor Brady's degrees are in mathematics (BSc and MSc from Manchester University, and PhD from the Australian National University). At Manchester University, he was awarded the Renold Prize as the outstanding undergraduate of his year. Professor Brady combines his work at Oxford University, where he founded the Robotics Laboratory and the Medical Vision Laboratory (MVL), with a range of entrepreneurial activities. He is Director of the EPSRC/MRC Inter-disciplinary research consortium on From Medical Images and Signals to Clinical Information (The IRC involves Oxford University, UCL, King's College London (including Guy's Hospital), and the University of Manchester. It also involves clinicians from over 30 specialities and is supported by over 20 companies. The chair of the Industrial Advisory Group will be Professor Will Stewart FREng, who is also a member of the Advisory Panel of the Department of Engineering Science.) He was appointed Senior Research Scientist of the MIT Artificial Intelligence Laboratory in 1980, and helped found its world famous robotics laboratory. In 1985, he left MIT to take up a newly created Professorship in Information Engineering. Professor Brady serves as a non-executive director and Deputy Chairman of Oxford Instruments plc, as a non-executive director of AEA Technology, and, until recently, Isis Innovation (Oxford University's intellectual property company). Professor Brady is a founding Director of the start-up companies Guidance and Control Systems and Mirada Solutions.

Professor Brady is the author of over 275 articles in computer vision, robotics, medical image analysis, and artificial intelligence, and the author or editor of nine books, including: Robot Motion (MIT Press, 1984), Robotics Science (MIT Press, 1989), Robotics Research (MIT Press 1984), and Mammographic Image Analysis (Kluwer, January 1999). He was Editor of the Artificial Intelligence Journal (1987–2002), and founding Editor of the International Journal of Robotics Research (1981–2000). He is a member of the Editorial Board of fourteen journals, most recently Medical Image Analysis.

Professor Brady was elected a Fellow of the Royal Academy of Engineering (UK) in 1991 and a Fellow of the Royal Society (UK) in 1997. He is a Fellow of the Institution of Electrical Engineers and a founding Fellow of the Association of Artificial Intelligence, and a Fellow of the Institute of Physics. He is a member of the Conseil Scientifique de l'INRIA France. He has been awarded honorary doctorates by the universities of Essex, Manchester, Liverpool, Southampton, and Paul Sabatier (Toulouse). He was awarded the IEE Faraday Medal for 2000 and the IEEE Third Millennium Medal for the UK.

α Scale spaces filters for phase based edge detection in Ultrasound images

α SCALE SPACES FILTERS FOR PHASE BASED EDGE DETECTION IN ULTRASOUND IMAGES

Ahror Belaid* Djamal Boukerroui*

*Université Abderrahmane Mira de Béjaïa, Laboratoire d'Informatique Médicale (LIMED),
06000, Béjaïa, Algérie

*Université de Technologie de Compiègne - CNRS, Heudiasyc UMR 7253 BP 20529 -
60205 Compiègne Cedex, France

ABSTRACT

An inherent characteristic of coherent imaging, including ultrasound imaging, is the presence of speckle noise with strong inhomogeneities. It makes modelling difficult and therefore it complicates their processing. However, it has been previously shown that these images respond well to local phase-based methods. Furthermore, the recent re-emergence of the α scale spaces theory opens new possibilities of phase-based image processing. In this paper, we make use of this new unified representation to derive new families of bandpass quadrature filters. This construction leads to a generalised α kernel filters including the commonly known families derived from the Gaussian and the Poisson kernels. The properties of each family are first presented and then, experiments on realistic simulations of ultrasound images are shown to demonstrate how the suggested filters can be used for edge detection.

Index Terms— α scale spaces, Quadrature filters, Edge detection, Local phase information, Ultrasound images.

1. INTRODUCTION

Ultrasound B-mode images are known to have low signal-to-noise ratio, low contrast, and high amounts of speckle, which is a correlated and multiplicative noise that inherently occurs in all types of coherent imaging systems. Moreover, ultrasound data often presents missing boundaries of the object of interest due to problems of specular reflection, shadows, signal dropout and attenuation. Consequently, ultrasound image segmentation is strongly influenced by the quality of data [1]. The literature on the subject is very abundant and interestingly edge based approaches are well represented in this literature. Possibly because, as it has been pointed in [1], that echocardiography has been one of the driving application areas of medical ultrasound and the most popular approach to endocardial segmentation has been to treat it as a contour finding problem. Thus, in this study we focus on boundary detection on ultrasound images.

The extensive literature on the subject suggests that conventional intensity gradient-based methods have had limited success on typical clinical images [1], mainly because of low signal-to-noise ratio and also because of attenuation. Solutions using local phase information are successfully used in [2, 3, 4]. Indeed, phase based methods are theoretically intensity invariant and work better on ultrasound images.

The estimation of the local signal properties (local phase, amplitude and orientation) are based on the calculation of the analytical representation of the signal [5, 6]. However, a calculation of these local quantities cannot be done directly in a phase-based technique. In practice, it uses a pair of bandpass quadrature filters (an even filter and odd filter) [7]. What is quite certain is that the estimation is intrinsically noisy and depends critically on the choice of the quadrature filters pairs [8]. Furthermore, as these local properties are scale dependent, their use for feature detection requires also their scale invariance (at least in a certain range) in order to detect only salient features and not noise. Therefore, the only reasonable approach that has proven itself is to combine several scales. As mentioned earlier, this is in favour for the development of scale space representations and theories. It is in this double context of scale space and quadrature filters that Felsberg et al. [9, 10] have introduced the Monogenic Poisson scale space. This new theory opened new possibilities of phase-based image processing in scale space.

Recently, Duits et al. [11] worked on a generalised form of scale space filters, appeared initially in [12], named the α scale spaces. It is a parameterised class ($\alpha \in]0; 1[$) of linear scale space representations which allows a continuous connection between the Poisson scale space ($\alpha = 1/2$) and the well known Gaussian scale space ($\alpha = 1$). In this paper, we make use of this new unified α scale space representation to derive new families of bandpass quadrature filters. These filters are built from derivatives and difference of the α scale space generating kernel. Thereby, they lead to the commonly known families of filters: Difference/Derivative of Poisson [10] and Gaussian filters. Following [8], the properties of each family are studied in terms of bandwidth, tuning frequency and

unit normalisation constant. These definitions are presented in section 2. In section 3, we investigate the use of these filters for edge detection using the Monogenic Feature Asymmetry (FAM) measures [3]. A comparison with the Log-Gabor filter and preliminary results on simulated ultrasound images are also presented. Section 4 provides a discussion followed by some concluding remarks in section 5.

2. NEW BAND-PASS QUADRATURE FILTERS

Given a 1D real signal $f(x)$, its α scale spaces representation $v_s^{(\alpha)} : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ is given by means of a convolution operation [11]:

$$v^{(\alpha)}(x, s) = \left(K_s^{(\alpha)} * f \right) (x) , \quad (1)$$

where s represents the scale parameter and $\alpha \in]0, 1]$. The kernel $K^{(\alpha)}$ is defined in the Fourier domain by the following expression [11]:

$$\mathcal{K}^{(\alpha)}(\omega, s) = \exp(-s|\omega|^{2\alpha}) . \quad (2)$$

In the following section, we aim to derive two new quadrature families using the above generating kernel of the α scale space.

2.1. α scale spaces derivative filters

Following the same procedures as reported by Boukerroui et al. [8], we propose a new generalised formulation of a whole parameterised class of bandpass filters. This new family of filters is based on the derivatives of the generating kernel of the α scale space and leads to the Poisson derivatives (PoD) and the Gaussian Derivatives (GD) families. Thus, we define the following family of 1D α Scale Spaces Derivative (ASSD) quadrature filters in the frequency domain as follow:

$$\mathcal{F}_{ASSD}(\omega) = \begin{cases} n_c \omega^a \exp(-(s\omega)^{2\alpha}) & \text{if } \omega \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where the derivative parameter $a \in \mathbb{R}^+$, meaning we are using fractional order derivatives. In order for the filters to satisfy the DC condition and to be also invariant to an additive ramp, we impose $a > 1$. n_c is a normalisation constant [8]. Important properties and tuning parameters of the ASSD filter, namely, the tuning frequency, filter's normalisations and the bandwidth, are given in the following propositions.

Proposition 2.1 *The peak tuning frequency of the scale space derivative filters is given by:*

$$\omega_0 = \frac{1}{s} \left(\frac{a}{2\alpha} \right)^{\frac{1}{2\alpha}} . \quad (4)$$

Proposition 2.2 *The unit normalisation constant n_c of the scale space derivative filters is given by:*

$$n_c = 2 \frac{\sqrt{\pi\alpha} 2^{\frac{2\alpha+1}{4\alpha}} s^{a+\frac{1}{2}}}{\sqrt{\Gamma\left(\frac{2\alpha+1}{2\alpha}\right)}} . \quad (5)$$

Proposition 2.3 *The octave bandwidth of the scale space derivative filters is given by:*

$$\beta = \frac{\ln\left(\frac{\mathcal{W}(-1, \mu)}{\mathcal{W}(0, \mu)}\right)}{2\alpha \ln(2)} , \mu = -\frac{1}{e^{\frac{2\alpha}{a}}} . \quad (6)$$

where $\mathcal{W}(k, \cdot)$ is the k^{th} branch of the Lambert function and $k = 0$ or -1 .

2.2. Difference of α scale spaces filters

A second way to build a bandpass filter given a low pass filter, is to use the difference operator. Thus, we study a new generalised bandpass filter build from the Difference of two α Scale Spaces filters (DoSS). The impulse response of this new filter, in the frequency domain, is given by :

$$\mathcal{F}_{DoSS}(\omega) = n_c \exp\left(-\frac{1}{2}(s_1\omega)^{2\alpha}\right) - n_c \exp\left(-\frac{1}{2}(s_2\omega)^{2\alpha}\right) , \quad (7)$$

where $\omega \geq 0$ and $s_1 < s_2$.

Proposition 2.4 *The peak tuning frequency of the difference of scale space filters is given by:*

$$\omega_0 = \frac{(4\alpha)^{\frac{1}{2\alpha}}}{s_2} \left(\frac{\log(\gamma)}{\gamma^{2\alpha} - 1} \right)^{\frac{1}{2\alpha}} \text{ with } \gamma = \frac{s_1}{s_2} . \quad (8)$$

Proposition 2.5 *The difference of scale space filters unit normalisation constant is given by:*

$$n_c = \frac{2\sqrt{\pi\alpha s_2}}{\sqrt{\Gamma\left(\frac{1}{2\alpha}\right)}} \left[1 + \frac{1}{\gamma} - 2 \frac{2^{\frac{1}{2\alpha}}}{s_2^{2\alpha-1}(\gamma^{2\alpha} + 1)} \right]^{-\frac{1}{2}} . \quad (9)$$

Note that the bandwidth of the DoSS filter is evaluated numerically; see Fig. 1. Furthermore, observe that there is a common value, $\beta = 2.5$, that defines the ASSD filter's bandwidth almost for all values of α . This will be useful to set a common value for β during the experiments over all possible values of α .

3. APPLICATION TO EDGE DETECTION IN ULTRASOUND IMAGES

Our aim in this section is to investigate the behaviour of the proposed filters on contour detection. Specifically, we are

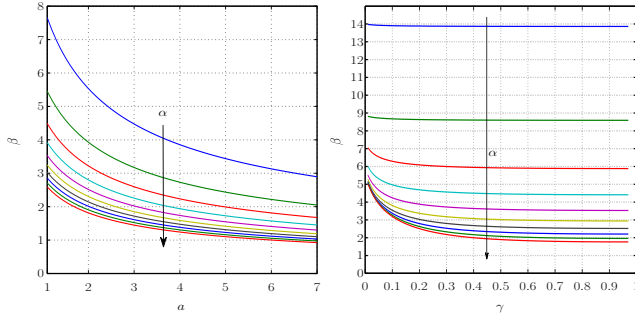


Fig. 1. Bandwidth of the ASSD filter as function of the derivative parameter a (left) and the DoSS filter as function of the scales ratio γ (right) for $\alpha \in \{0.1, 0.2, \dots, 1\}$.

interested in the influence of the parameter α on the detection's performance. To this end, we use the *Monogenic Feature Asymmetry* (FA_M) as a measure of contour detection. It is a phase based measure and uses multiple scales for the analysis of the given image.

3.1. The FA_M measures

The monogenic signal at a given scale s of a 2D signal $f(\mathbf{x})$ can be represented by a scalar valued even and a vector valued odd filtered responses, with the following simple tick:

$$\begin{aligned} \text{even}_s &= c * f, \\ \text{odd}_s &= (c * h_1 * f, c * h_2 * f), \end{aligned}$$

where c is the spatial domain representation of an isotropic bandpass filter tuned at scale s , and $\mathbf{h}(\mathbf{x}) = (h_1, h_2)(\mathbf{x}) = \frac{1}{2\pi} \frac{\mathbf{x}}{|\mathbf{x}|^3}$ is the generalised 2D Hilbert transform kernel [9, 10]. We define the multiple scales monogenic Feature Asymmetry by [3]:

$$FA_M = \frac{1}{N} \sum_s \frac{||\text{odd}_s| - |\text{even}_s| - t_s|}{\sqrt{\text{even}_s^2 + |\text{odd}_s|^2 + \varepsilon}}, \quad (10)$$

where N is the total number of scales, $[\cdot]$ denotes zeroing of negative values and t_s is a noise threshold estimated similarly as in [6].

3.2. Evaluation and results

The FA_M is evaluated on 120 realistic simulated ultrasound images with known ground truth. More details on these data can be found in [13], to which the reader is referred to. Some typical images are shown in Fig. 3. In our experimental study, the bandwidth β is set to $\{2.5, 3.6\}$ respectively for the ASSD and DoSS filters. 2.5 octaves is the only value possible for all $\alpha \in [0, 1]$, and 3.6 octaves is the lowest possible value for $\alpha \in [0.5, 1]$. We found by experiment that the wavelength of the finest scale $w = 12$ pixels was appropriate for most of our data. In order to detect fine

as well as coarse structures, we consider three scales where the scaling factor between successive filters is set to 2.1.

The evaluation is carried out by comparing machine generated contours to ground-truth data using the *precision-recall* framework. In our context, the precision-recall curves are obtained by varying the detection threshold. There is an interesting point on this curve defined by the F measure, given as

$$F = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

Thus, the location of the maximum of the above measure along the curve defines the optimal threshold and provides a summary score. Herefore, following [14], we use the below quantities as evaluation scores: the best F cost on the dataset for a fixed scale (ODS)¹, the average F cost on the dataset for the best scale per image (OIS), and finally the average precision (AP) on the full recall range (equivalently, the area under the precision-recall curve).

Table 1. Results for FA_M contour detector using the ASSD, the DoSS and the Log-Gabor (LG) filters. Shown are the optimal value of α corresponding to the F-measure scores when choosing an Optimal Dataset Scale (ODS) or an Optimal Image Scale (OIS), as well as the Average Precision (AP). 120 simulated US images were used.

Filter	Parameter	ODS	OIS	AP
ASSD	$\alpha = 0.2$	0.69	0.71	0.54
DoSS	$\alpha = 0.9$	0.66	0.69	0.54
LG	$\beta = 3.6$	0.65	0.68	0.52

4. RESULTS AND DISCUSSION

Table 1 reports the main results of the contour detection experiments. More details, as well as the precision-recall curves of the results are shown in Fig. 2 (right). First, we notice that in general the results obtained by using the ASSD family outperforms significantly both the DoSS and the LG filters, and this for optimal values of α . Indeed, Fig. 2 shows that in a wide range, approximately $\alpha \in [0, 0.85]$, the ASSD filters outperforms clearly the DoSS filters. While beyond $\alpha = 0.85$, the DoSS filters outperforms slightly the ASSD one. One reason for this may be the fact that the DoSS family have a larger bandwidth, since the lowest possible value in the range of $\alpha \in [0.5, 1]$ is 3.6 octaves. Therefore, it is more sensitive to noise and performs less on textured images. This observation is also supported by the few illustrative synthetic images shown in figures Fig. 3

Secondly, although the experimental study presented here is limited, it suggests that the ASSD filter with $\alpha = 0.2$ is more efficient ($F=0.693$) than the commonly used Gaussian

¹Here we mean by scale, the detection threshold.

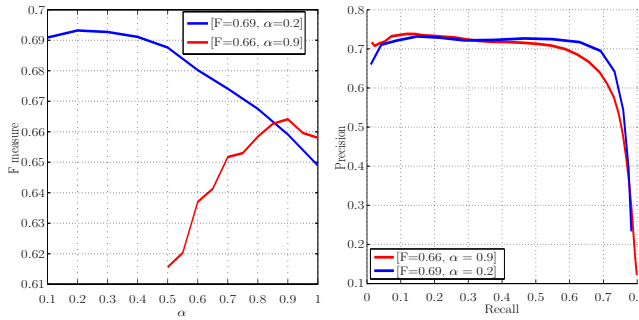


Fig. 2. F-measure as function of the shape parameter α (left) and Precision-Recall curves (right). The highest F-measure and the corresponding value of α are also reported for ASSD (blue line) and DoSS (red line) filters.

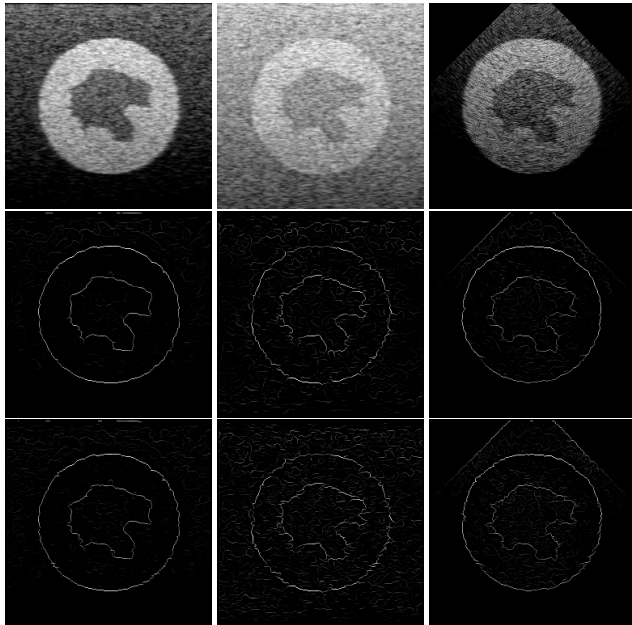


Fig. 3. Illustrative F_A edge detection results using the ASSD and the DoSS filters on simulated US images with different: tissues characteristics, attenuation level and log compression parameters. From top to bottom: original images, edge detection using ASSD ($\alpha = 0.2$) and DoSS ($\alpha = 0.9$) filters respectively.

Derivative ($\alpha = 1$, $F=0.648$) and Poisson Derivative ($\alpha = 1/2$, $F=0.687$). Also, the DoSS filter with $\alpha = 0.9$ is more efficient ($F=0.664$) than Difference of Gaussian ($F=0.658$) and Difference of Poisson filters ($F=0.615$). Finally, the performances of the Log-Gabor filter seem to be close to the DoSS filter. This is confirmed by the close F measures in Table 1.

5. CONCLUSION

Two new parameterised classes ($\alpha \in]0; 1[$) of band-pass quadrature filters are presented. We then looked more closely

at the influence of the shape parameter α in the context of phase based edge detection on ultrasound images. Based on a quantitative evaluation on 120 simulated images, the preliminary results show that the proposed ASSD and DoSS filters with specific values of α outperform the commonly used Log-Gabor and the special cases Derivative/Difference of Gaussian ($\alpha = 1$) and Poisson ($\alpha = 1/2$) filters.

6. REFERENCES

- [1] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, 2006.
- [2] M. Mulet-Parada and J. A. Noble, "2D+T acoustic boundary detection in echocardiography," *Med. Image Anal.*, vol. 4, no. 1, pp. 21–30, 2000.
- [3] A. Belaid, D. Boukerroui, Y. Maingourd, and J-F. Lerallut, "Phase based level set segmentation of ultrasound images," *IEEE Trans. Info. Tech. Biomed.*, vol. 15, no. 1, pp. 138–147, 2011.
- [4] A. Belaid, D. Boukerroui, Y. Maingourd, and J-F. Lerallut, "Implicit active contours for ultrasound images segmentation driven by phase information and local maximum likelihood," in *ISBI*, Chicago Illinois, USA, Apr. 2011, pp. 630–635.
- [5] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognit. Lett.*, vol. 6, no. 5, pp. 303–313, 1987.
- [6] P. Kovess, "Image features from phase congruency," *Videre: J. Comput. Vis. Res.*, vol. 1, no. 3, pp. 1–26, 1999.
- [7] H. Knutsson and M. Andersson, "Whats so good about quadrature filters," in *IEEE Int. Conf. Image Proc.*, 2003, pp. 61–64.
- [8] D. Boukerroui, J. A. Noble, and M. Brady, "On the choice of band-pass quadrature filters," *J. Math. Imag. Vision*, vol. 21, no. 1, pp. 53–80, 2004.
- [9] Michael Felsberg and Gerald Sommer, "Scale adaptive filtering derived from the laplace equation," in *Proc. of the 23rd DAGM-Symposium on Pattern Recognition*, London, UK, 2001, pp. 124–131, Springer-Verlag.
- [10] M. Felsberg and G. Sommer, "The monogenic scale-space: A unifying approach to phase-based image processing in scale-space," *J. Math. Imag. Vision*, vol. 21, no. 1, pp. 5–26, 2004.
- [11] Remco Duits, Michael Felsberg, Luc Florack, and Bram Platel, " α scale spaces on a bounded domain," in *Proc. Int. Conf. on Scale Space Methods in Computer Vision*, Berlin, Heidelberg, 2003, pp. 494–510, Springer-Verlag.
- [12] E.J. Pauwels, L.J. Van Gool, P. Fiddelaers, and T. Moons, "An extended class of scale-invariant and recursive scale space filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 7, pp. 691–701, 1995.
- [13] D. Boukerroui, "A local rayleigh model with spatial scale selection for ultrasound image segmentation," in *Proc. of British Machine Vision Conf.*, Surrey, UK, Sep 2012, pp. 84–84.
- [14] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, 2011.

Enhancement of contrast regions in suboptimal ultrasound images with application to echocardiography



● *Original Contribution*

ENHANCEMENT OF CONTRAST REGIONS IN SUBOPTIMAL ULTRASOUND IMAGES WITH APPLICATION TO ECHOCARDIOGRAPHY

DJAMAL BOUKERROU,*, J. ALISON NOBLE,* MARC C. ROBINI** and MICHAEL BRADY*

*Medical Vision Laboratory, Department of Engineering Science, University of Oxford, Parks Road,
Oxford OX1 3PJ, UK

**CREATIS, CNRS Research Unit (UMR 5515), INSA de Lyon, FRANCE

(Received 6 June 2001; in final form 2 October 2001)

Abstract—In this paper we propose a novel feature-based contrast enhancement approach to enhance the quality of noisy ultrasound (US) images. Our approach uses a phase-based feature detection algorithm, followed by sparse surface interpolation and subsequent nonlinear postprocessing. We first exploited the intensity-invariant property of phase-based acoustic feature detection to select a set of relevant image features in the data. Then, an approximation to the low-frequency components of the sparse set of selected features was obtained using a fast surface interpolation algorithm. Finally, a nonlinear postprocessing step was applied. Results of applying the method to echocardiographic sequences (2-D + T) are presented. The results demonstrate that the method can successfully enhance the intensity of the interesting features in the image. Better balanced contrasted images are obtained, which is important and useful both for manual processing and assessment by a clinician, and for computer analysis of the sequence. An evaluation protocol is proposed in the case of echocardiographic data and quantitative results are presented. We show that the correction is consistent over time and does not introduce any temporal artefacts. (E-mail: djamal@robots.ox.ac.uk) © 2002 World Federation for Ultrasound in Medicine & Biology.

Key Words: Image enhancement, Feature detection, Image quality, Phase congruency, Surface interpolation, Echocardiography, Ultrasound.

INTRODUCTION

The first step toward automatic analysis or evaluation of images frequently consists of feature detection and segmentation. In the field of medical image analysis, robustness, accuracy and reproducibility are critical. Traditionally, in early (low-level) processing and analysis methods, the data are used directly without feature enhancement or bias field (*i.e.*, intensity inhomogeneity) correction. The most important preprocessing step investigated and applied is denoising, which improves the signal-to-noise ratio (SNR). However, denoising is only a partial solution because it cannot correct the image for artefacts introduced by the imaging system. Indeed, it often depends on the imaged object, which finally leads to nonhomogeneous regions in the image. Intensity inhomogeneities are a well-studied problem in the area of

image analysis of magnetic resonance imaging (MRI) and ultrasound (US) B-scan images (Hughes and Duck 1997; Xiao et al. 2000). Recently, following the publication of the work by Wells et al. (1996) on bias field correction of MRI data, several authors have investigated this problem and several approaches have been proposed for MRI (Guillemaud and Brady 1997; Velthuisen et al. 1998). Although bias field correction is often necessary for good segmentation, many approaches have exploited the idea that a good segmentation also helps estimation of the bias field (Ahmed et al. 1999; Leemput et al. 1999; Zhang et al. 2001). By contrast, intensity inhomogeneity correction for US images has received relatively little attention, possibly because of the high noise level of B-mode images. Some recent intensity-based adaptive segmentation approaches, which intrinsically take into account the nonuniformity of the tissue classes, have yielded promising results (Ashton and Parker 1995; Boukerroui et al. 1999; Boukerroui 2000). More recently, a novel technique for finding acoustic boundaries

Address correspondence to: Dr. Djamal Boukerroui, Old Library, Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK. djamal@robots.ox.ac.uk

in 2-D and 2-D + T echogram sequences has been proposed (Mulet-Parada 2000; Mulet-Parada and Noble 2000). The most important advantage of this technique is its intensity-independence. However, because the noise rejection in this method involves an intensity-based noise threshold, the method is not truly intensity-invariant and is highly sensitive to noise. The 2-D + T version of the published technique takes advantage of temporal continuity to improve its robustness to noise and detect only relevant and continuous features over time. The authors have reported that spatiotemporal estimation is insufficient for low frame-rate sequences and that there are a number of localisation problems because of the nonuniformity of wall velocity during the cardiac cycle (Mulet-Parada 2000). This underlines the need for development of a feature enhancement approach to correct the image. To our knowledge, the first attempt to adapt bias field correction of Wells et al. (1996) to B-scan US data is the method proposed in (Xiao et al. 2000). Results shown for breast and cardiac US images demonstrate that it can successfully remove intensity inhomogeneities, and significant improvement is achieved in tissue contrast and in the resulting image segmentation. The approach is promising. However, as reported in (Xiao et al. 2000), it still requires user interaction to set the image model parameters.

In this paper, we propose a novel feature-based contrast enhancement approach. Our approach is based on a phase-based feature detection algorithm, followed by sparse surface interpolation and subsequent nonlinear postprocessing. We first exploited the intensity-invariant property of phase-based acoustic feature detection to select the relevant features in the data. Then, an approximation to the low-frequency components of the sparse set of selected features was obtained, using a surface interpolation algorithm. Finally, a nonlinear postprocessing step with one control parameter was applied.

The paper is organised as follows. In the Methods section, we briefly describe the mathematical framework underlying the sparse surface interpolation algorithm (Yaou and Chang 1994), review the feature asymmetry (FA) measure for 2-D acoustic boundary detection (Mulet-Parada 2000), and then present the nonlinear postprocessing stage. The proposed evaluation protocol in the special case of echocardiographic data, as well as quantitative results using the approach, are presented and discussed in the Results section. The authors' conclusions are summarised in the Conclusion section.

METHODS

Overview of algorithm

A block diagram of our new feature-based contrast enhancement method is shown in Fig. 1. First, features are detected in the image $I(x, y)$. This provides a nor-

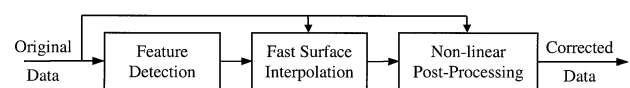


Fig. 1. Block diagram of the feature enhancement method.

malised likelihood image $FA_{2-D}(x, y)$ where the intensity value at a position (x, y) is proportional to the significance of the detected features. The $FA_{2-D}(x, y)$ measure varies from a maximum of 1 (indicating a very significant feature) down to 0 (indicating no significance). The feature detector that we used is based on phase congruency (PC) (Morrone and Owens 1987; Kovess 1999) because it provides a single unified theory that is capable of detecting a wide range of features, rather than being specialised for a single feature type, such as an intensity step edge. Inspection of cardiac B-mode US images reveals features of a number of quite distinctive kinds; for example, the intensity ridge corresponding to (a large part of) the epicardial wall and the mitral valve. Further, PC is theoretically invariant to brightness and contrast. Hence it is, in principle, robust against typical variations in image formation. The disadvantage of PC is a direct result of contrast invariance; namely, its high sensitivity to noise. The poor SNR of B-mode US, including heavy speckle, means that this is a problem that has particularly to be addressed when applying PC to cardiac US.

Following feature detection, the sparse data at feature locations are then interpolated by a fast sparse surface interpolation algorithm using the likelihoods to estimate the degradation field; then, a novel nonlinear processing method using the degradation field is applied to the original data to enhance or de-emphasise feature values. These steps are described next.

2-D sparse surface interpolation

In this section, we review the method we employ for fast sparse surface interpolation. This is based on Yaou and Chang (1994), to which the reader is referred for further details.

Surface interpolation from a sparse set of noisy measured points is an ill-posed problem because an infinite set of surfaces can satisfy any given set of constraints. Hence, a regularisation procedure, taking into account the visual relevance and computational efficiency is usually applied, so that the interpolation problem becomes a minimisation of an energy functional of the form:

$$U(f) = U_d(f, d) + \lambda U_r(f). \quad (1)$$

The first term in eqn (1) is a faithfulness measure of the approximating surface f to the measured data d and is called the "cost or constraint functional." The second is the regularisation functional; λ is a nonnegative (Lagrange multi-

plier) parameter controlling the amount to which the data are to be considered (piecewise) smooth. A commonly used cost functional is the weighted sum of squares:

$$U_d(f, d) = \sum_{i=1}^n w_i (f(x_i, y_i) - d_i)^2, \quad (2)$$

which measures the difference between the measured field $d = \{(x_i, y_i, d_i)\}$ and the approximating surface $f(x_i, y_i)$; and where $w = \{0 \leq w_i \leq 1\}$ is the corresponding set of weights for the measured field, reflecting the confidence of the measured information at each position ($w_i = 0$ means the absence of information at (x_i, y_i)); n being the number of data points. Regarding the regularisation term, a common approach has been to use a variational functional to constrain the solution, often expressed as a thin plate energy, otherwise known as the quadratic variation (Grimson, 1981):

$$U_r(f) = \iint \left[\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right] dx dy. \quad (3)$$

In general, obtaining an analytic solution of the Euler–Lagrange equations resulting from the above optimisation problem is difficult. Therefore, an approximation to the continuous problem using discrete operators is used. This leads to a numerical solution. Suppose that the data d is defined on a regular rectangular lattice $G = \{(x_i, y_j), 1 \leq i, j \leq N\}$, and that a discrete representation of the surface is defined using a set of nodal variables $\mathbf{v} = \{v_{i,j} = f(x_i, y_j)\}$. The discrete representation of the cost functional in eqn (2) is:

$$U_d(\mathbf{v}, d) = \sum_{i,j} w_{i,j} (v_{i,j} - d_{i,j})^2. \quad (4)$$

By concatenating all the nodal variables $v_{i,j}$ and the data $d_{i,j}$, respectively, into column vectors \mathbf{v} and \mathbf{d} , we obtain the usual matrix representation of eqn (4),

$$U_d(\mathbf{v}, d) = (\mathbf{v} - \mathbf{d})^T \mathbf{A}_w (\mathbf{v} - \mathbf{d}), \quad \mathbf{A}_w = \text{diag}(\{w_{i,j}\}). \quad (5)$$

Regarding the regularisation term, the finite element method provides a continuous surface approximation, which is a good means of converting the continuous expression for the energy and leads to a tractable discrete problem that has a numerical solution. The discrete form of the thin plate is given by:

$$U_r(\mathbf{v}) = \sum_{i,j} \left\{ \begin{aligned} &(\mathbf{v}_{i+1,j} - 2\mathbf{v}_{i,j} + \mathbf{v}_{i-1,j})^2 \\ &+ 2(\mathbf{v}_{i+1,j+1} - \mathbf{v}_{i,j+1} - \mathbf{v}_{i+1,j} + \mathbf{v}_{i,j})^2 \\ &+ (\mathbf{v}_{i,j+1} - 2\mathbf{v}_{i,j} + \mathbf{v}_{i,j-1})^2 \end{aligned} \right\}, \quad (6)$$

which can be reorganised as follows:

$$U_r(\mathbf{v}) = \sum_{i,j} \sum_{m,n} v_{i,j} a_{i,j,m,n} v_{m,n}, \quad (7)$$

where the coefficients $a_{i,j,m,n}$ describe the relations between the nodal variables $v_{i,j}$ and $v_{m,n}$ and are called relation coefficients. Hence, the discrete version of the continuous regularisation term given by eqn (3) is:

$$U_r(\mathbf{v}) \propto \mathbf{v}^T (\lambda \mathbf{A}_r) \mathbf{v}, \quad (8)$$

where \mathbf{A}_r is $N^2 \times N^2$ sparse matrix and contains at most 13 nonzero elements per line. Finally, by adding eqns (5) and (8), we obtain the corresponding discrete version of the functional in eqn (1):

$$U(\mathbf{v}) = \mathbf{v}^T \mathbf{A} \mathbf{v} - 2\mathbf{v}^T \mathbf{b} + c, \quad (9)$$

where, $\mathbf{A} = \lambda \mathbf{A}_r + \mathbf{A}_w$ and $\mathbf{b} = \mathbf{A}_w \mathbf{d}$, and c is a constant. The resulting energy function has a minimum at $\mathbf{v} = \mathbf{v}^*$, the solution of the linear system $\mathbf{A} \mathbf{v} = \mathbf{b}$, with a very sparse system matrix \mathbf{A} .

Resolution in the wavelets space

The surface interpolation problem described in the last section leads to the solution of a large linear system with a sparse system matrix \mathbf{A} . Therefore, the equation system is nearly singular and results in poor convergence when simple iterative methods are used. To obtain fast surface interpolation, a scheme is needed that can improve the numerical conditioning. The multigrid (Terzopoulos 1983) and hierarchical basis techniques (Szeliski 1990) have been applied successfully to speed up convergence. Both of these techniques use a multiresolution approach. However, in both cases, the spatial frequency domain property of the interpolation problem is not effectively utilized. Recently, a more tractable approach, in terms of simplicity and efficiency, has been proposed (Yaou and Chang 1994). The approach utilizes the concept of preconditioning in a wavelet transform space. Preconditioning techniques aim to reduce the condition number κ of the system matrix (defined as the ratio between the magnitude of the largest and the smallest eigenvalues), which implies speeding up the rate of convergence. Wavelet transform analysis provides a powerful way to simultaneously filter data in space and scale and is easily implemented using filter banks (Mallat

1999). Yaou and Chang (1994) use vector space decomposition reconstruction of the wavelet transform to analyse the multiresolution representation of the surface interpolation problem. The authors exploit the fact that, usually, the high-frequency component of the interpolation problem converges much faster than the low-frequency component in surface interpolation. Interestingly, this is the same observation that inspired Terzopoulos' use of the multigrid algorithm. The use of the discrete wavelet transform (DWT) allows the low-frequency component and the high-frequency component of the interpolation problem to be solved separately.

In other words, the minimisation is carried out in a wavelet space using an asynchronous iterative computation and a biorthogonal spline wavelet basis for the preconditioning step (Yaou and Chang 1994). The DWT preconditioning transfers the equation system to an equivalent one with new nodal variables $\tilde{\mathbf{v}}$ and a new system matrix, $\tilde{\mathbf{A}}$, that is much denser than the original one, \mathbf{A} . This implies that a more global connection between the interpolation nodes can be made which considerably improves the convergence rate of the iterative solution. To give the reader some insight on the relative convergence rate with and without preconditioning in our problem, for a 64×64 interpolation problem, the value of κ is reduced by a factor of 2.5 and 5.4, respectively, when two and three decomposition levels for the DWT are used. We refer the reader to Yaou and Chang (1994) for more details about fast surface interpolation using the multiresolution wavelet transform, and to Cohen et al. (1992), Daubechies (1992), and Mallat (1999) regarding the biorthogonal wavelet basis and the associated filter banks.

Phase-based feature detection

Phase-based feature detection has been investigated extensively following the publication of the local energy model of feature detection (Morrone and Owens 1987). This model postulates that features can be defined and classified using their phase signatures, or their phase congruency (PC). In other words, feature information is encoded at points where all the Fourier components are maximally in phase. Thus, the PC is maximum. These observations have led to the development of a number of phase-based feature detection algorithms (Kovesi 1999; Mulet-Parada 2000; Mulet-Parada and Noble 2000) (and references therein). In particular, measures based on phase information seem to be more appropriate for acoustic feature detection because US images are characterized by a low SNR due to the presence of speckle and to the high range of imaging artefacts, causing the alteration of the intensity magnitude of equally significant features in the data.

Strictly speaking, the concept of PC is only defined

in one dimension because its definition involves the Hilbert transform. Typically, the computation of PC (and the related concept of local energy) uses a pair of quadrature filters (Venkatesh and Owens 1990), normally log-Gabor filters. A series of oriented 2-D filters can, for example, be constructed by 'spreading' a log-Gabor function into 2-D. In this way, extension to 2-D of the 1-D phase measure may be obtained (Kovesi 1999).

In our work, we have used the 2-D feature asymmetry (FA) measure used in Mulet-Parada (2000), and Mulet-Parada and Noble (2000) for feature detection. This measure provides good detection of asymmetric image features, such as step edges, and has the advantage of being intensity invariant. The 2-D FA measure is defined by:

$$FA_{2-D}(x, y) = \sum_m \frac{\lfloor |o_m(x, y)| - |e_m(x, y)| - T_m \rfloor}{\sqrt{o_m^2(x, y) + e_m^2(x, y) + \varepsilon}}, \quad (10)$$

which is a sum over m orientations of a normalised measure of the difference between the odd $o_m(x, y)$ and the even $e_m(x, y)$ filter responses. Here, $\lfloor \cdot \rfloor$ denotes zeroing of negative values, ε is a small positive number to avoid division by zero and T_m is an orientation-dependent noise threshold, defined by (Mulet-Parada 2000):

$$T_m = k \cdot std[|o_m(x, y)| - |e_m(x, y)|], \quad (11)$$

where k is a positive factor controlling the noise threshold and std is the standard deviation (SD). For more details about the implementation and the setting parameters of the log-Gabor filters and the spreading functions, see Kovesi (1999), Mulet-Parada (2000) and Mulet-Parada and Noble (2000).

The output of applying the FA algorithm is a location image, which is a normalised likelihood image $[0, 1]$, where the intensity values can be interpreted as a confidence measure of feature detection.

The new feature-enhancement algorithm

Having reviewed the two principal algorithms that are used in our new feature detection method, we now address how these are used for feature enhancement. Briefly, our method involves reconstructing an approximation to the intensity inhomogeneities that can be utilised to correct the original corrupted data. A mathematical model for the intensity inhomogeneity in US images was developed by Hughes and Duck (1997). The authors used a multiplicative degradation model. Motivated by this, we define a correction equation as:

$$I_c(x, y) = \frac{I(x, y)/\max(I(x, y))}{\nu^*(x, y)/\max(\nu^*(x, y)) + \gamma}. \quad (12)$$

Here, $\nu^*(x, y)$ is the interpolated surface and γ is a positive control parameter that ensures that $I_c(x, y) \propto I(x, y)$ for $\gamma \gg 1$. The maximum correction is obtained when $\gamma \ll 1$. Assuming that the image intensities of occurrences of a single tissue type should be equal, an estimate of the low frequency components of an intensity data field can be made by taking the image intensity values of the image only at the locations of the relevant features. An estimate of the base frequency of this degradation can be found using the fast surface interpolation algorithm as follows. We define the set of nodal variables ν and the corresponding weighting field w , by:

$$\nu = \{\nu_{i,j} = \max_{B_{i,j}} I(x, y) \text{ if } FA_{2-D}(x_i, y_j) > 0; \quad 1 \leq i, j \leq N\} \quad (13)$$

$$w = \{w_{i,j} = FA_{2-D}(x_i, y_j); 1 \leq i, j \leq N\}. \quad (14)$$

Here, $B_{i,j}$ is a small rectangular window centred at pixel position (x_p, y_j) and N is the size of the data. Taking the maximum intensity value in a window centred on the feature position guarantees that we always take the highest value of the step edge. Hence, eqns 13 and 14 define the two inputs for the surface interpolation stage, which provide us with an estimate of the low-frequency image degradation.

Implementation details

In this section, we present some implementation details related to our method and its efficiency. The surface interpolation problem is difficult because the nodal variables are defined on a very sparse grid. For instance, in the particular case of echocardiography images, the nonexistence of any constraint variables (nodes) in the left ventricular cavity causes problems at the end diastolic period of the cardiac cycle when the area of the cavity is very large. On the other hand, derivation of the system matrix \tilde{A} requires $O(n^2)$ operations where n is the number of the nodal variables, so we aim to keep n small. To avoid these two problems, an estimation of the surface using low-pass filtered and subsampled data is carried out; then we interpolate the resulting surface on a regular grid to provide the estimate of the bias field of the original data. We found that the use of low-pass downsampled data benefits both the fast surface interpolation stage and the 2-D FA detection stage, because the filtering operation allows reduction of the speckle noise. Finally, a histogram threshold is found to eliminate the high values due to the enhancement of

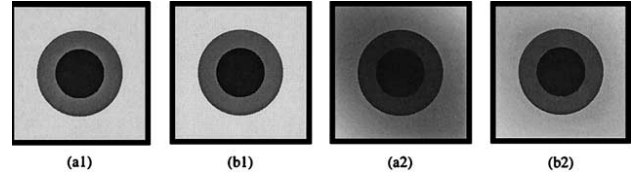


Fig. 2. Original images: (a1) Ideal; (a2) corrupted, and their corresponding enhanced images (b1) and (b2). Results obtained for $N = 64$ and $\gamma = 0.2$.

noise peaks. The threshold h , is defined empirically by the following equation:

$$P_{I_c}(I_c(x, y) < h) = 0.999. \quad (15)$$

Here, $P_{I_c}(\cdot)$ is a probability measure. The corrected image is rescaled between [0 to 255] and the grey value of the rejected points is set to 255 (see Fig. 2c1). We found the rejection of a small number of points sufficient to cut the tail of the distribution without introducing false edges or modifying the shape of the distribution. Cutting the tail of the distribution (which contains no information) has the advantage of enlarging the intensity range of the corrected data.

Note that the nonlinear processing given by eqn (12) and the thresholding operation are applied to the whole sequence when the processed data are 2-D + T.

RESULTS AND QUANTITATIVE EVALUATION

All the results presented in this paper are obtained for $k = 0.45$ (FA threshold factor), $\lambda = 1$ (FSI parameter) and with 2 decomposition levels for the DWT in the fast surface interpolation algorithm.

Synthetic data

The aim of the first experiment was to show, for synthetic data, that the proposed approach is capable of removing (or at least reducing) the bias field without introducing any artefacts. Figure 2 shows two images, an ideal one (a1) and corrupted one (a2) and their corrections (b1 and b2, respectively). We can see that a significant contrast enhancement is obtained in both cases and that the corrected images appear visually similar.

Real data

One short-axis and two long-axis sequences of approximately 100 frames were chosen randomly from the ADEQUATE project database of routine acquisition scans at the Hospital General Universitario Gregorio Marañón, Madrid and at the John Radcliffe Hospital, Oxford. The data sets were acquired on an HP-SONOS

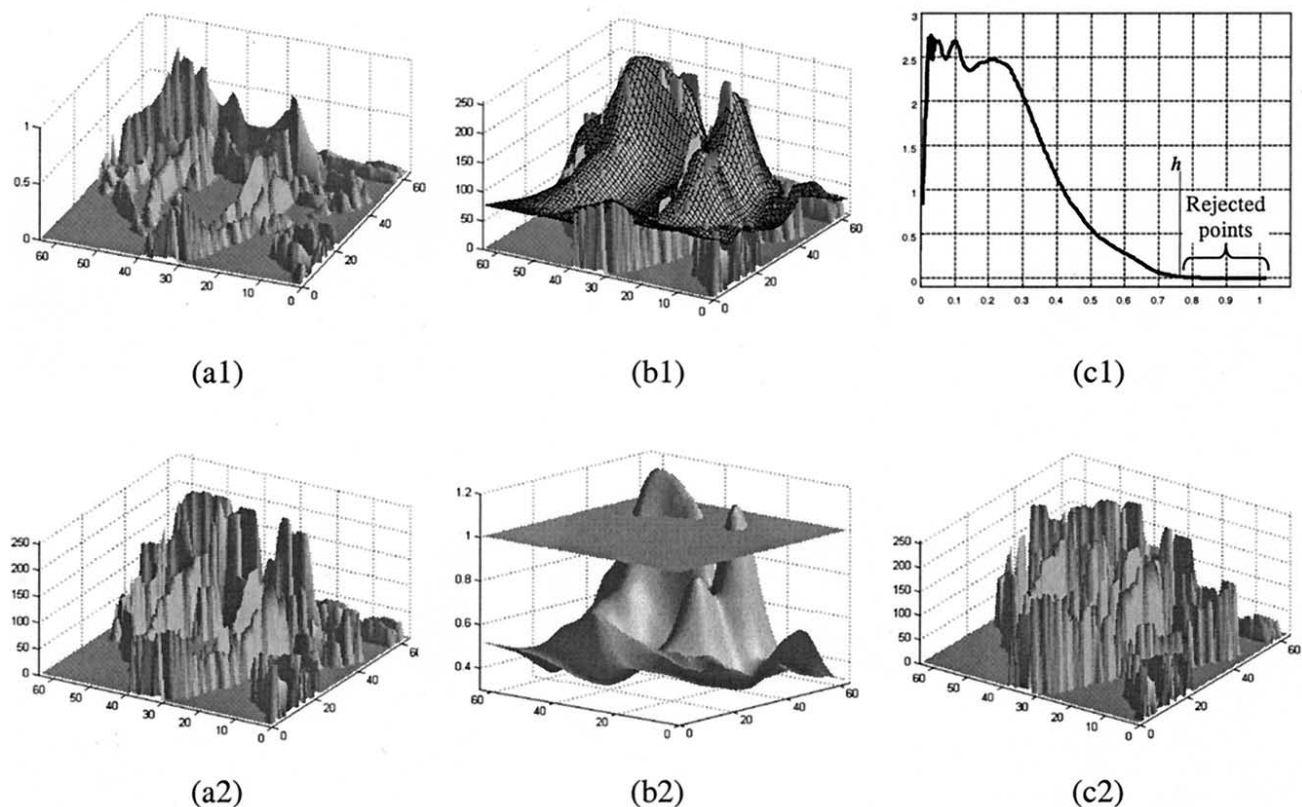


Fig. 3. Illustration of steps in the algorithm. (a1) Likelihood image representing the weighting field w , for the surface interpolation stage; (a2) the original data at the location of the detected boundaries representing the data field v , for the surface interpolation stage; (b1) the result of surface interpolation superimposed on the data field; (b2) the normalised interpolated surface with the additional shift, $\gamma = 0.2$, showing clearly the region where the intensity will be lowered and the region where it will be enhanced; (c1) histogram thresholding operation; (c2) corresponds to (a2), but shows it after correction.

5500 scanner (Agilent Technologies). Data sets 1 and 3 were fundamental mode acquired at a frame rate of approximately 80 Hz and transducer frequency of 3.6 MHz. Data set 2 was acquired at a frame rate of 25 Hz, using second harmonic image enhancement (transmit/receive 1.8/3.6 MHz). The size of the original images was 720×512 pixels. A region-of-interest (ROI) of 512×512 in the centre of each frame was extracted.

Figure 3 shows the original data and results at intermediate stages of processing for an echocardiographic image. The image used in this experiment is shown in Fig. 4a1. Notice the correlation between the likelihood image (a1) and the intensity image of the detected features (a2). Hence, if the SNR is low, the 2-D FA measure does not yield a clean feature detection image. In that case, either the noise threshold, eqn (11) has to be set to a higher value, which increases the false-negative detection rate, or it has to be set to a low value, in which case the false-positive rate will increase. Comparison of the images (a2) and (c2) provides some (qualitative) insight about how much the features have been enhanced.

To understand the effect of using approximate data at a coarse resolution, we compared feature enhancements using different sampling resolutions. As illustrated by the example shown in Fig. 4, we have found that we do not observe any significant difference between the results obtained using approximations to the original data (512×512) on a 64×64 or a 128×128 grid. Both corrections seem to give acceptable corrections and contrast enhancement (see the profile lines). However, the plot of the vertical line profile Fig. 4c3 shows that the background signal (interior of the LV cavity) is more enhanced for $N = 128$ than for $N = 64$. This is due to the problem of sparse surface interpolation mentioned in the Implementation Details section. Figure 4b1, b2, b3 show the results of the 2-D FA boundary detection on the original image and on the enhanced images. Significant improvements in boundary detection on the enhanced images is observed, particularly in the apex region where the intensity values of the original image are very low.

In the third experiment, we studied the influence of the control parameter γ , eqn (12). Figure 5 shows the

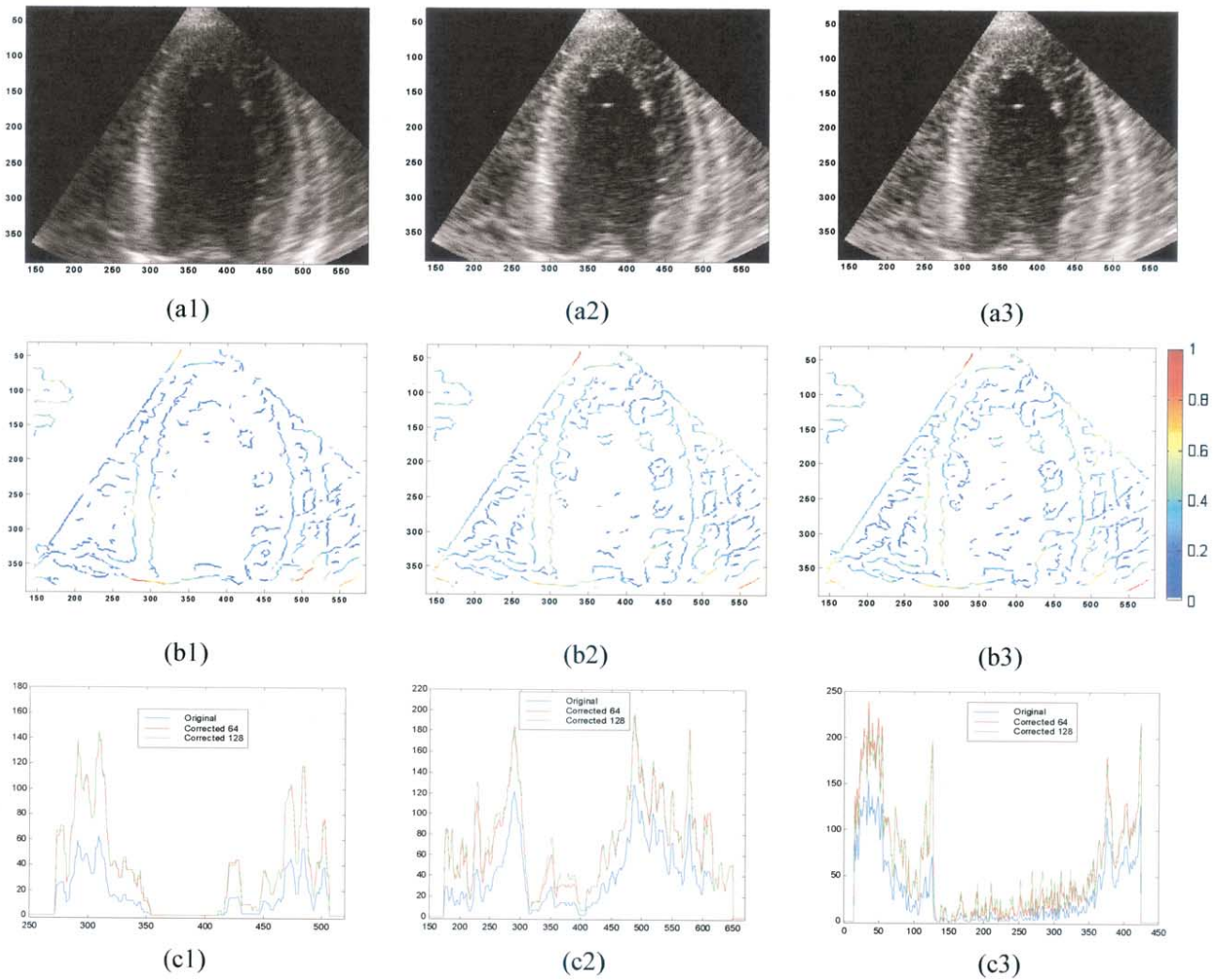


Fig. 4. Frame 11 of data set 1. Comparison of (a1) the original image and (a2) the enhanced one obtained with $N = 64$ and (a3) $N = 128$. Images (b) show the corresponding feature asymmetry results (likelihood image). Figures (c) show profile lines: (c1) horizontal line 150; (c2) horizontal line 350; and (c3) vertical line 355. Results obtained with $\gamma = 0.2$.

enhancement results and the corresponding feature asymmetry detection on the corrected image for three different values of γ : 0.4, 0.2 and 0.05. The three results are good and appear, at first sight, to be similar. However, on more careful observation, we note that the enhancement increases when γ decreases. The plots of line profiles shown in Fig. 6 offer a clear demonstration of the influence of γ . Notice that the three results appear to be in good agreement where the signal is high and low signal values are more enhanced for $\gamma = 0.05$ than 0.2 or 0.4. However, this observation does not mean that the enhancement result for $\gamma = 0.05$ (or $\gamma < 0.05$) is better than for the other two. Indeed, if $\gamma = 0.05$ enhances the low signal values better than $\gamma = 0.2$ (0.4), it does the opposite for high signal values. An objective evaluation and quantitative results of the enhancement are necessary to answer the question as to which values of γ give the

best enhancement. This is described next. Figure 7 shows the enhancement results and the corresponding feature asymmetry detection for the original and the corrected images for the two other data sets (second harmonic long axis and B-mode short axis).

Quantitative evaluation. A standard protocol in the evaluation of heart disease is to assess regional and global left ventricular function, using quantitative and qualitative measurement based on the myocardial walls (epicardial and endocardial). This includes heart wall delineation, wall thickness estimation and motion estimation, like muscle contraction and the opening and closing of the cardiac valves. Clinicians desire better image quality. The automatic analysis of echocardiograms is, to date, a challenge due to the low contrast and very noisy appearance of standard grey-level (B-mode) images. The quantitative evalua-

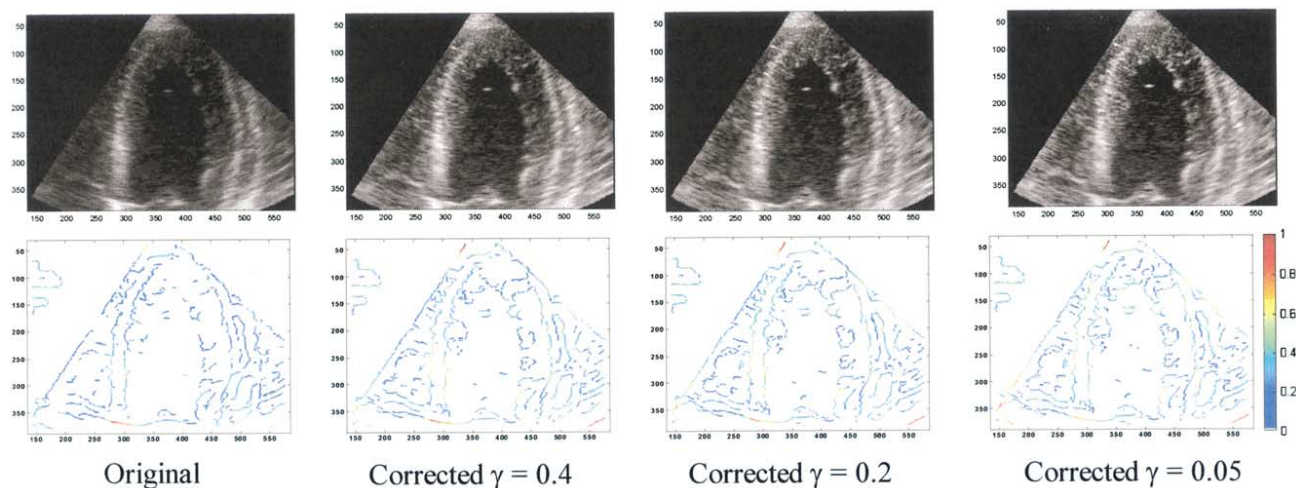


Fig. 5. Influence of the control parameter γ . The original images correspond to frame 13 of data set 1. Results obtained for $N = 64$.

tion of the proposed feature-based contrast enhancement approach will be described next.

Quantitative evaluation of computer vision algorithms is an important task, particularly in the case of medical imaging. The availability of ground truth makes this task easier. However, there is no ground truth for the data available for the current study. To overcome this problem, we have to answer the question: Which are the important features in the data and how can an automatic detection algorithm respond to the features and not others? In the case of echocardiographic images, and for the purpose of this paper, we are interested in the detection of the endocardial boundary. Because these features are often modelled as step edges, a measure of the height of the step is a good evaluation parameter. A second criterion that our enhancement should satisfy is consistency over time (*i.e.*, the method should not introduce temporal

artefacts). A good measure of temporal consistency is the correlation coefficient between measures calculated on the enhanced images and the same measures calculated on the original data. We investigate both aspects next.

For each image, we define three regions next to the interesting features: RC is the region located in the cardiac cavity near to the endocardial wall, RM is the myocardium, and RE is the region located from the epicardial border outwards. Briefly, for each image, seven segments were defined by selecting points on a manually defined interpolating spline fit to the endocardial border. The intensity profiles crossing the three regions (RE, RM, RC) could then be automatically defined by computing the normal to the curve and sampling pixels along the normal direction, as defined before. The width of each region was about five pixels. This is illustrated in Fig. 8. For each of the seven segments, the

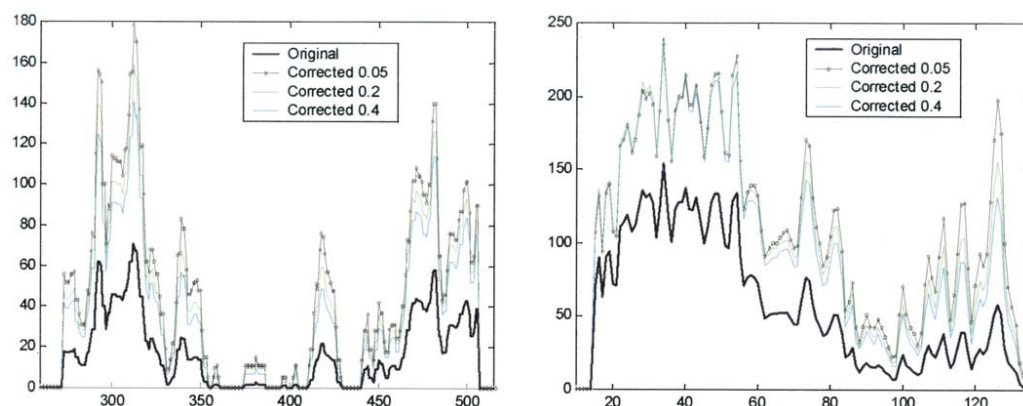


Fig. 6. Influence of γ . Line profiles of the results shown in Fig. 5. Left: Horizontal line 150; right: vertical line 355. Observe, on the right side of the figure, the enhancement of the peak corresponding to the endocardium border on the apex.

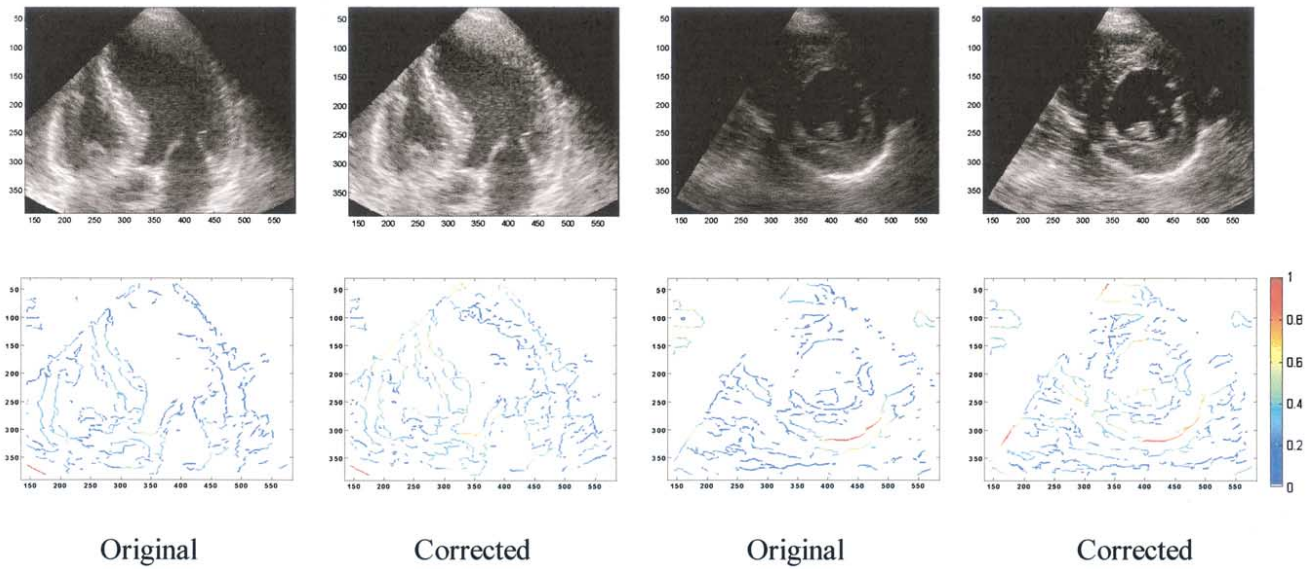


Fig. 7. Correction results on two other data sets. Left: data set 2, frame 17; right: data set 3, frame 0. Results obtained for $N = 64$ and $\gamma = 0.2$.

contour was sampled at 600 points to generate 600 intensity profiles. In practice, we tried to avoid the papillary muscles in the cavity. In the case of the short-axis images, it was difficult to know where the epicardial border was, so we did not measure RE.

For each image, we computed ten measures: mean (signal) and SD of the cavity signal (RC), of myocardium signal (RM), of the signal beyond the epicardial border

(RE) and of the differences (RE - RM and RM - RC). We then computed the mean and the SD (μ_t , σ_t) of each of these measures over time. From this, we could investigate if there was a significant contrast improvement, especially at features location, namely the cardiac walls (see RE - RM and RM - RC values).

Table 1 shows the computed measures on a long axis data set for the original images and those enhanced for $N = 64$ and $N = 128$. Note the similarity of the calculated measures and the small relative differences, demonstrating the nonexistence of significant difference between the two results. This table confirms the conclusion of experiment 2 because the highest relative difference is for the calculated measures for the cardiac cavity region (RC) (the cavity signal is more enhanced for $N = 128$).

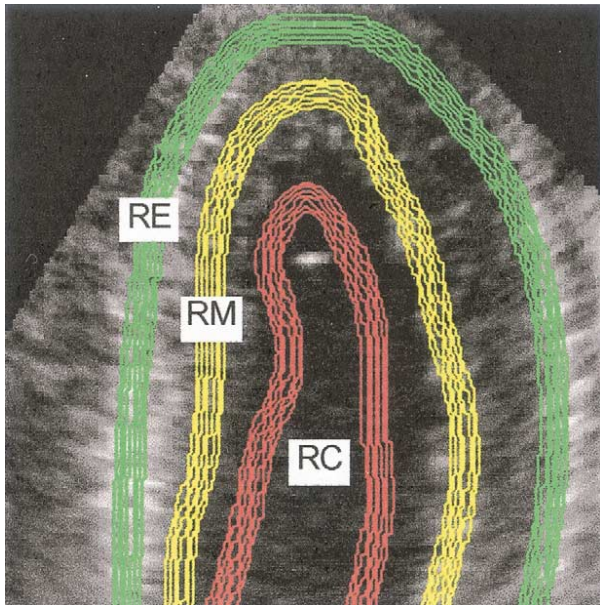


Fig. 8. Illustration of region-of-interest for the evaluation protocol.

Table 1. Quantitative comparison of the enhancement for two different sampling resolutions (64 and 128)

Data set		Original	Corrected	Corrected 128
1 Long axis		(μ_t , σ_t)	64 (μ_t , σ_t)	(μ_t , σ_t)
RC	Signal	6.81, 1.02	16.37, 2.16	18.88, 2.57
	SD	6.94, 1.14	13.07, 2.22	15.59, 2.65
RM	Signal	29.70, 3.29	61.86, 5.45	64.01, 5.47
	SD	18.61, 2.21	29.62, 2.42	30.38, 2.38
RE	Signal	61.18, 3.99	107.64, 5.38	107.07, 5.33
	SD	35.53, 3.20	44.71, 3.59	45.14, 3.87
RE - RM	Signal	31.48, 4.12	45.78, 6.10	43.06, 6.33
	SD	28.59, 2.63	44.96, 3.12	44.88, 3.11
RM - RC	Signal	22.89, 3.20	45.49, 5.70	45.14, 5.64
	SD	15.98, 1.77	28.34, 2.35	29.13, 2.37

Here, signal is the spatial mean of the signal. The table shows the means and the SDs over the frames of one cardiac cycle. $\gamma = 0.2$.

Table 2. Quantitative analysis of the enhancement and study of the influence of the control parameter γ on the enhancement for the three data sets

		Original (μ_r, σ_r)	Corrected 0.4 (μ_r, σ_r)	Corrected 0.2 (μ_r, σ_r)	Corrected 0.1 (μ_r, σ_r)	Corrected 0.05 (μ_r, σ_r)
Data set 1 Long axis						
RC	Signal	6.81, 1.02	14.80, 1.94	16.37, 2.16	17.65, 2.32	18.49, 2.42
	SD	6.94, 1.14	12.17, 2.02	13.07, 2.22	13.59, 2.36	13.90, 2.44
RM	Signal	29.70, 3.29	56.16, 5.23	61.86, 5.45	66.09, 5.51	68.84, 5.48
	SD	18.61, 2.21	28.51, 2.64	29.62, 2.42	30.29, 2.19	30.86, 2.07
RE	Signal	61.18, 3.99	101.71, 5.33	107.64, 5.38	110.89, 5.43	112.40, 5.47
	SD	35.53, 3.20	45.93, 3.65	44.71, 3.59	42.98, 3.53	41.59, 3.49
RE-RM	Signal	31.48, 4.12	45.54, 5.84	45.78, 6.10	44.81, 6.31	43.56, 6.48
	SD	28.59, 2.63	42.67, 3.15	44.96, 3.12	46.72, 3.13	48.03, 3.23
RM-RC	Signal	22.89, 3.20	41.36, 5.33	45.49, 5.70	48.43, 5.85	50.35, 5.90
	SD	15.98, 1.77	26.20, 2.26	28.34, 2.35	30.18, 2.52	31.66, 2.74
Data set 2 Long axis						
RC	Signal	6.45, 0.64	13.88, 1.12	15.25, 1.20	16.39, 1.27	17.11, 1.33
	SD	6.63, 0.67	11.20, 0.97	11.89, 1.05	12.26, 1.12	12.46, 1.19
RM	Signal	30.57, 3.54	57.52, 5.64	63.17, 5.92	67.24, 6.05	69.81, 6.11
	SD	19.16, 2.18	29.36, 2.87	30.39, 2.80	30.84, 2.68	31.09, 2.59
RE	Signal	61.87, 3.78	102.56, 4.89	108.39, 4.91	111.52, 4.99	112.90, 5.10
	SD	35.67, 2.49	46.18, 3.11	45.03, 3.23	43.35, 3.32	42.01, 3.40
RE-RM	Signal	31.29, 4.63	45.04, 6.50	45.22, 6.75	44.27, 6.97	43.09, 7.14
	SD	28.19, 2.35	42.03, 2.64	44.24, 2.70	45.86, 2.82	46.99, 3.01
RM-RC	Signal	24.12, 3.49	43.63, 5.63	47.92, 5.98	50.86, 6.18	52.69, 6.29
	SD	16.42, 2.03	26.48, 2.84	28.29, 2.85	29.66, 2.84	30.64, 2.86
Data set 3 Short axis						
RC	Signal	1.41, 0.62	4.31, 1.64	5.05, 1.95	5.68, 2.14	5.53, 2.09
	SD	2.05, 0.79	5.59, 1.65	6.56, 1.94	7.36, 2.14	7.18, 2.10
RM	Signal	18.55, 3.40	43.17, 8.90	50.76, 11.12	56.28, 12.95	54.78, 12.99
	SD	12.04, 1.21	24.94, 2.11	28.85, 2.61	31.72, 3.25	30.72, 3.47
RM-RC	Signal	17.14, 3.01	38.86, 7.74	45.71, 9.72	50.60, 11.39	49.25, 11.46
	SD	12.11, 1.33	25.14, 2.13	29.02, 2.30	31.86, 2.57	30.80, 2.63

Here, signal is the spatial mean of the signal. The table shows the means and the SDs over the frames. $N = 64$.

Table 2 presents the evaluation measures for the third experiment (influence of γ). Note the small values of the SD over time for all the computed measures, both for the original and the enhanced images, indicating a good time consistency. We observe that the signal enhancement is proportional to γ^{-1} and the SDs for RC and RM increase slightly. However, the RE SD decreases with γ^{-1} . This is because, as we have already noted, the highest grey-level values will be reduced, and the lower grey-level intensities will be increased (see Fig. 3b2). When the enhancement is high (γ small), a “saturation phenomenon” appears at the highest intensity values. Because the RE region corresponds to the highest grey values in the images, the spatial SD of this region will decrease with enhancement.

Analysis of the signal differences reveals that both the signal and the SD of the difference $RM - RC$ increase as γ decreases. However, the signal of the difference $RE - RM$ increases as γ decreases until $\gamma = 0.2$ and decreases for $\gamma > 0.1$. This is a consequence of the “saturation phenomenon” as the step edge $RE - RM$ is at high intensities.

These observations enable us to understand more fully the behaviour of the enhancement in function of the

parameter γ . The reader may notice how close the quantitative measures are for the different values of the control parameter. In our experiments, we found that a value between 0.1 and 0.2 gives acceptable enhancement results and good trade-offs between the enhancement of the low ($RM - RC$) and high ($RE - RM$) grey-level intensity features.

Consistency. Because two key parts of our feature enhancement algorithm do not take into account the temporal information, the consistency of the enhancement over time needs to be studied to show that no temporal artefacts are introduced. Figure 9 shows the spatial correlation curves (where correlation has been computed between consecutive frames) for the original sequences and for their corresponding enhanced sequences. In other words, each curve shows the temporal evolution of the spatial correlation coefficient of consecutive frames in the sequence. The interesting aspect of these curves is not the absolute values of the spatial correlation, but its evolution over the frames. If the enhancement method is consistent over time, the evolution of the curves over time should be the same. Indeed, the curves show that the temporal correlation of the

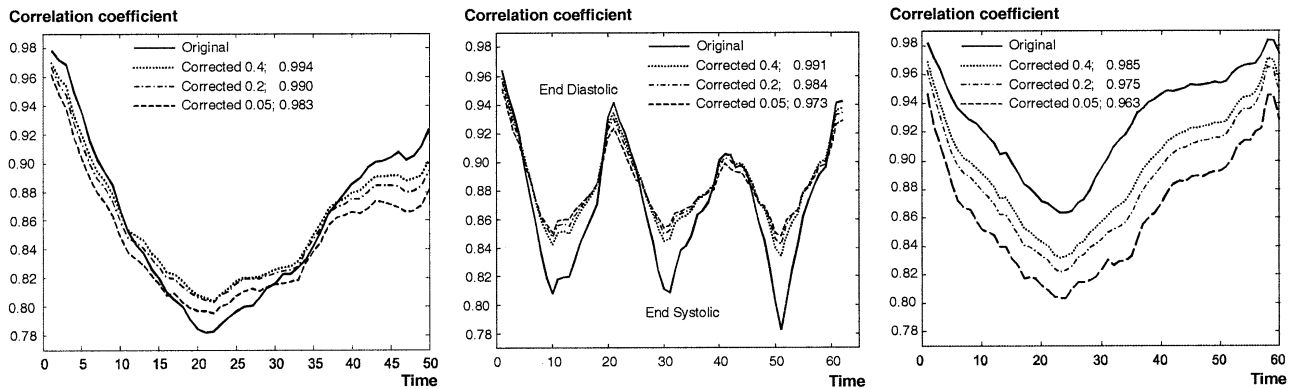


Fig. 9. Evolution over time of the spatial correlation coefficients between consecutive frames. Left: Data set 1; middle: data set 2; right: data set 3. “Corrected 0.2; 0.990” means results obtained for $\gamma = 0.2$ and the correlation coefficient of the curve to the equivalent curve of the original data is 0.990. $N = 64$.

original data is well conserved in the corrected sequences (see the temporal correlation coefficients in the figure legend). Notice that the temporal continuity is better conserved for the two first data sets (long axis) than for the third one (short axis). The lowest correlation coefficient for the temporal continuity is 0.963 obtained for the enhancement of data set 3 with $\gamma = 0.05$.

The second experiment analysing temporal consistency involved computing the spatial correlation coefficient between the enhanced and the original data frame-by-frame. If the enhancement method is consistent over time, the value of the spatial correlation should be almost constant for all the frames. In other words, we expect the correlation coefficient to be independent of time. Table 3 shows the mean and the SD over the time of this measure for the three processed data sets. Notice how small the SDs are, indicating good temporal consistency of the enhancement.

A comparison between the results of the three data sets gives us some insight into the robustness of the proposed approach.

Finally, Fig. 10 shows further enhancement results obtained on four different data sets (one frame each). The grey-level data were acquired in the same condition as “before” data. All examples are typical of those in clinical practice (*i.e.*, were not specifically acquired for this

paper, but were routine scans). The last example (right image) was acquired using a contrast enhanced power Doppler (CEPD) imaging system. CEPD detects the contrast agent in the cardiac cavity that has been previously injected into the patient. The enhancement results suggest that the proposed method could be used as a pre-processing for CEPD images because it reduces the intensity inhomogeneity of the signal in the cavity region. This potentially suggests that our method may be very useful for improving automatic detection of the endocardial border. This is currently being investigated in the ADEQUATE project.

CONCLUSION

The performance of the proposed feature enhancement has been illustrated on two test images and on several echocardiographic sequences. An evaluation protocol has been proposed in the case of echocardiographic images and quantitative results have been presented. The proposed feature enhancement has one control parameter, γ ; its influence on the results has also been studied. The consistency over time of the enhancement of the proposed approach has been demonstrated to ensure that no temporal artefacts are introduced. This is an important point, both for manual processing and assessment by a clinician, and for computer analysis of the sequence. The corrected images produced by the new method facilitate visual diagnosis by a clinician because the contrast between the heart wall and the cavity is enhanced. On the other hand, significant improvement in the results of the 2-D FA detection algorithm has been noted, in comparison with its application on the nonenhanced data.

The approach is original and gives good enhancement results on echocardiographic image sequences. The performance of the proposed approach is even more convincing if the processed data are displayed as a movie

Table 3. Evaluation experiment of the consistency over time

	$\gamma = 0.4; (\mu_r, \sigma_r)$	$\gamma = 0.2; (\mu_r, \sigma_r)$	$\gamma = 0.05; (\mu_r, \sigma_r)$
Data set 1	0.966, 0.003	0.947, 0.004	0.914, 0.007
Data set 2	0.980, 0.003	0.967, 0.005	0.942, 0.009
Data set 3	0.938, 0.003	0.906, 0.004	0.849, 0.010

This table shows the mean and the SD over the frames of the spatial correlation coefficient between the corrected images to their corresponding original ones. $N = 64$.

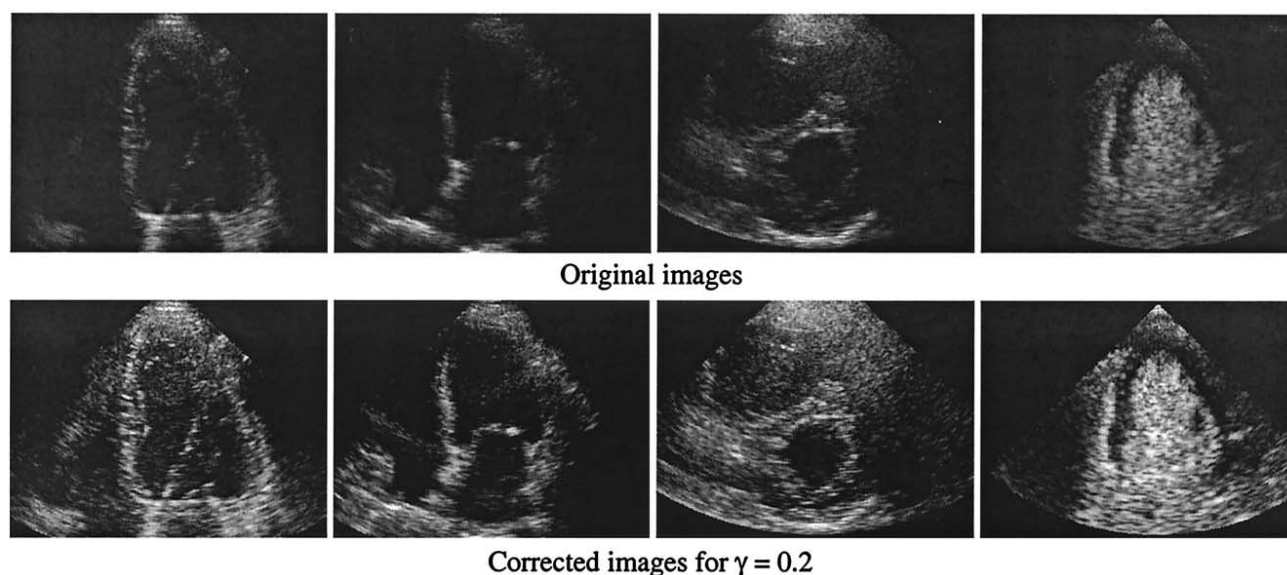


Fig. 10. Original and enhanced images of four different echocardiographic studies, $\gamma = 0.2$, $N = 64$. From left to right: B-mode long axis, B-mode four-chamber, B-mode short axis, CEPD long axis.

sequence.¹ The proposed approach is relatively fast in comparison with the existing algorithms in the literature (13 s/image 512×512 on a Pentium 400 MHz), but would require further speed-up for real-time processing and its use in a clinical environment (currently it takes 21 min to process one study of 100 frames).

Acknowledgements—The authors are grateful to Dr. M. M. Parada and Dr. J. Declerck, from Mirada Solutions Ltd, Oxford, for providing software used in part of this work. We also thank N. Clarke, from John Radcliffe Hospital, Oxford, and J. Bermejo, from Hospital General Universitario Gregorio Marañón, Madrid. This work was supported by the EC-funded ADEQUATE project (IST: 1999-10837).

REFERENCES

- Ahmed MN, Yamany SM, Mohamed NA, Farag AA. A modified fuzzy c-means algorithm for MRI bias estimation and adaptive segmentation. Cambridge, UK: MICCAI, 1999:72–81.
- Ashton EA, Parker KJ. Multiple resolution Bayesian segmentation of ultrasound images. *Ultrason Imaging* 1995;17:291–304.
- Boukerroui D. Segmentation bayésienne d'images par une approche markovienne multiresolution. Application à l'imagerie échocardiographique et échographique 2D et 3D du sein. Ph.D. thesis. CRE-ATIS, INSA 502. Villeurbanne Cedex (France), 2000.
- Boukerroui D, Basset O, Baskurt A, Noble A. Segmentation of echocardiographic data. Multiresolution 2D and 3D algorithm based on gray level statistics. Cambridge, UK: MICCAI, 1999:516–524.
- Cohen A, Daubechies I, Feauveau J-C. Biorthogonal bases of compactly supported wavelets. *Commun Pure Appl Math* 1992;45:485–560.
- Daubechies I. Ten Lectures on Wavelets. CBMS-NSF Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992:61.
- Grimson WEL. From images to surfaces: Computational study of the human early visual system. Cambridge, MA: MIT Press, 1981.
- Guillemaud R, Brady M. Estimating the bias field in MR images. *IEEE Trans Med Imaging* 1997;16:238–251.
- Hughes DI, Duck FA. Automatic attenuation compensation for ultrasonic imaging. *Ultrasound Med Biol* 1997;23:651–664.
- Kovesi P. Image feature from phase congruency. *Videre: J Comp Vision Res* 1999;1:1–26.
- Leemput KV, Maes F, Vandermeulen D, Suetens P. Automated model-based bias field correction of MR images of the brain. *IEEE Trans Med Imaging* 1999;18:885–896.
- Mallat S. A wavelet tour of signal processing. New York: Academic Press, 1999.
- Morrone MC, Owens RA. Feature detection from local energy. *Pattern Recogn Lett* 1987;6:301–313.
- Mulet-Parada M. Intensity independent feature extraction and tracking in echocardiographic sequences. Ph.D. thesis. Dept. of Eng. Science, Oxford University, 2000.
- Mulet-Parada M, Noble JA. 2D+T acoustic boundary detection in echocardiography. *Med Image Anal* 2000;4:21–30.
- Szeliski R. Fast surface interpolation using hierarchical basis functions. *IEEE Trans Pattern Anal Machine Intell* 1990;12:513–528.
- Terzopoulos D. Multilevel computational process for visual surface reconstruction. *Comput Vision Graphics Image Proc* 1983;24:52–96.
- Velthuisen RP, Heine JJ, Cantor AB, Lin H, Fletcher LM, Clarke LP. Review and evaluation of MRI nonuniformity correction for brain tumor response measurements. *Med Phys* 1998;25:1655–1666.
- Venkatesh S, Owens R. On the classification of image features. *Pattern Recogn Lett* 1990;11:339–349.
- Wells WM, Grimson EL, Kikinis R, Jolesz FA. Adaptive segmentation of MRI data. *IEEE Trans Med Imaging* 1996;15:429–442.
- Xiao G, Brady M, Noble JA, Zhang Y. Contrast enhancement and segmentation of ultrasound images—a statistical method. San Diego, USA: SPIE Med Imaging, Image Processing 2000; 3979: 1116–1125.
- Yao M-H, Chang W-T. Fast surface interpolation using multiresolution wavelet transform. *IEEE Trans Pattern Anal Machine Intell* 1994;16:673–688.
- Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 2001; 20:45–57.

¹ MPEG movies of the results are available at: <http://www.robots.ox.ac.uk/~djamal/umb/umb.html>

Velocity estimation in Ultrasound images : a block matching approach

Velocity estimation in Ultrasound images: a block matching approach

Djamal Boukerroui^{1,2}, J. Alison Noble² and Michael Brady²

¹ HEUDIASYC, UMR CNRS #6599, Université de Technologie de Compiègne,
BP 20529 - 60205 Compiègne Cedex, France.

djamal.boukerroui@hds.utc.fr

² Medical Vision Laboratory, Department of Engineering Science,
University of Oxford, Parks Road, Oxford OX1 3PJ, UK.

{noble,jmb}@robots.ox.ac.uk

Abstract. In this paper, we focus on velocity estimation in ultrasound images sequences. Ultrasound images present many difficulties in image processing because of the typically high level of noise found in them. Recently, Cohen and Dinstein have derived a new similarity measure, according to a simplified image formation model of ultrasound images, optimal in the maximum likelihood sense. This similarity measure is better for ultrasound images than others such as the sum-of-square differences or normalised cross-correlation because it takes into account the fact that the noise in an ultrasound image is multiplicative Rayleigh noise, and that displayed ultrasound images are log-compressed. In this work we investigate the use of this similarity measure in a block matching method. The underlying framework of the method is Singh's algorithm. New improvements are made both on the similarity measure and the Singh algorithm to provide better velocity estimates. A global optimisation scheme for algorithm parameter estimation is also proposed. We show that this optimisation makes an improvement of approximately 35% in comparison to the result obtained with the worst parameter set. Results on clinically acquired cardiac and breast ultrasound sequences, demonstrate the robustness of the method.

1 Introduction

The measurement of optical flow or image velocity is a fundamental problem in Computer Vision. Several techniques have been presented in the literature and many more continue to appear [1, 14, 7, 9, 2]. Such estimation or measurement of optical flow may be done, for example, to improve the efficiency of encoding the image, or to allow enhancement of the display of, or measurement of, the movement of some particular tracked part of the image to assist an observer to interpret the image [12]. Indeed, in medical applications for example, motion measurements constitute an essential component in the evaluation of any patient with known or suspected heart disease. In particular, detecting and characterising abnormalities in segmental wall motion function has become the hallmark of diagnosing coronary artery disease because reduced motion correlates with

2 Djamel Boukerroui et al.

ischaemic muscle action [3, 6]. Motion measurements are also used in breast deformation analysis to measure the elastic properties of tissues and to provide an indication of tissue hardness (cancerous tissue is, in general, harder than normal tissue) by computing a strain image or relative Young's modulus image [5].

In this paper, we focus on velocity estimation in ultrasound images sequences. Ultrasound images present many difficulties in image processing because of the typically high level of noise found in them. For example, the tracking of cardiac walls in cardiac ultrasound images is difficult because of the high level of noise and also because of the nature of the cardiac motion. Various ways of motion estimation in ultrasound sequences have been proposed [10–12, 4, 8, 9], but it is a difficult task in which there is room for improvement.

Optical flow methods can be classified as belonging to one of main three groups [1]: Differential techniques or gradient-based methods are based on the assumption that the brightness of a pattern is invariant over time and compute image velocity from spatio-temporal intensity derivatives and use a regularisation procedure based on *a priori* knowledge. Differential methods give good results on good quality images, however, they are highly sensitive to noise because of numerical differentiation and produce inaccurate results where the brightness constancy assumption is violated. In the second class, frequency-based techniques, two types of methods exist: energy-based and phase-based approaches. Theoretically, phase-based optical flow estimation is the most appropriated method for ultrasound images. The use of phase information makes the method robust to attenuation artefacts. The disadvantage of filter-based estimation is that the filter response is optimal only for a velocity range. We believe that is hard to design optimal filters tuned to the velocity range in a cardiac sequence as cardiac motion varies during the cardiac cycle. It has been reported in [11, 12] that spatio-temporal estimation is insufficient for low frame-rate sequences and that there are a number of localisation problems because of the non-uniformity of wall velocity during the cardiac cycle. The third class is block-matching motion estimation which defines velocity as the shift that yields the best fit between image regions/features at different times. The best match is found by maximising a similarity measure. Matching methods are in general computationally expensive for dense flow field estimation. They are particularly well suited if the estimation of the flow field is necessary only at some location (eg. only at the heart walls). Motivated by the results obtained in [10] and a recent similarity measure derived according to a simplified image formation model of ultrasound images [4], we have developed a new block-matching method. The underlying framework of the method is Singh's algorithm [13].

The outline of this paper is as follows. First a brief description of the similarity measure used and the underlying assumptions are presented in Section 2. Section 3 describes Singh optical flow estimation with a focus on the new changes made to improve the estimation. Implementation details and the global optimisation scheme for the parameters estimation are given in section 4. Section 5, presents results on clinically acquired cardiac and breast ultrasound sequences, and the summary and conclusion appear in Section 6.

2 Maximum Likelihood motion estimation [4]

We assume that two consecutive frames \mathbf{x} and \mathbf{y} are the realisation of two random variables \mathbf{X} and \mathbf{Y} . Let's suppose that a block $\mathbf{x}_i = \{x_{ij}, j = 1 \dots \Lambda\}$ in \mathbf{x} matches a block $\mathbf{y}_i = \{y_{ij}, j = 1 \dots \Lambda\}$ in \mathbf{y} and the displacement vector is denoted by $\mathbf{v}_i = [u_i, v_i]^T$. Here, the indice i is an index over all possible blocks and j is an index of pixels in the block. Given the above notations the maximum likelihood (ML) estimation of \mathbf{v}_i is given by [4]:

$$\mathbf{v}_i^{ML} = \arg \max_{\mathbf{v}_i} p(\mathbf{x}_i | \mathbf{y}_i, \mathbf{v}_i) . \quad (1)$$

where the conditional probability density function (pdf) depends in general on the noise model. Note that the above equation supposes implicitly the whiteness of the noise as the velocity estimation of a block i is independent of the remaining displacement field.

In ultrasound, when the speckle is fully developed, the noise is multiplicative and follows a Rayleigh pdf. If the noiseless value of the j^{th} pixel in region \mathbf{x}_i is denoted by s_{ij} and under the assumption of independent noise, the following model for the observed pixels in \mathbf{x} and \mathbf{y} stands [4]:

$$y_{ij} = \eta_{ij}^1 s_{ij} , \quad (2)$$

$$x_{ij} = \eta_{ij}^2 s_{ij} , \quad (3)$$

where η_{ij}^1 and η_{ij}^2 are two independent noise elements with a Rayleigh density function given by:

$$p_{\eta^1}(\eta^1) = \frac{\eta^1}{\alpha^2} \exp \left\{ -\frac{(\eta^1)^2}{2\alpha^2} \right\} , \quad \eta^1 > 0 , \quad (4)$$

$$p_{\eta^2}(\eta^2) = \frac{\eta^2}{\beta^2} \exp \left\{ -\frac{(\eta^2)^2}{2\beta^2} \right\} , \quad \eta^2 > 0 . \quad (5)$$

Given these two models, the following equation is obtained:

$$x_{ij} = \eta_{ij} y_{ij} , \quad (6)$$

where the noise term is defined as follow:

$$\eta_{ij} = \frac{\eta_{ij}^1}{\eta_{ij}^2} \quad \text{and} \quad p_{\eta}(\eta) = 2 \left(\frac{\alpha}{\beta} \right)^2 \frac{\eta}{\left(\eta^2 + \left(\frac{\alpha}{\beta} \right)^2 \right)^2} , \quad \eta > 0 . \quad (7)$$

Taking the natural logarithm of both sides of eq. 6 and denoting $\ln(x)$ by \tilde{x} we obtain the following model for displayed ultrasound images:

$$\tilde{x}_{ij} = \tilde{y}_{ij} + \tilde{\eta}_{ij} , \quad (8)$$

and the pdf of the additive noise $\tilde{\eta}_{ij}$ is given by:

$$p(\tilde{\eta}) = 2 \left(\frac{\alpha}{\beta} \right)^2 \frac{\exp(2\tilde{\eta})}{\left(\exp(2\tilde{\eta}) + \left(\frac{\alpha}{\beta} \right)^2 \right)^2} . \quad (9)$$

4 Djamel Boukerroui et al.

Cohen and Dinstein [4] supposed that the independent noise of the successive frames follow the same distribution (i.e. $\alpha = \beta$) which simplifies the above equation. Given the above model (eq. 8), the conditional pdf is given by:

$$p(\tilde{\mathbf{x}}_i | \tilde{\mathbf{y}}_i, \mathbf{v}_i) = \prod_{j=1}^A 2 \frac{\exp(2(\tilde{x}_{ij} - \tilde{y}_{ij}))}{(\exp(2(\tilde{x}_{ij} - \tilde{y}_{ij})) + 1)^2} . \quad (10)$$

Motion based on the above equation was denoted by CD_2^1 by Cohen and Dinstein [4]. Maximising the above pdf is equivalent to maximising the following objective function:

$$E_i^{CD_2}(\mathbf{v}_i) = \sum_{j=1}^A \{ \tilde{x}_{ij} - \tilde{y}_{ij} - \ln(\exp(2(\tilde{x}_{ij} - \tilde{y}_{ij})) + 1) \} . \quad (11)$$

This similarity measure is better for ultrasound images than others such as the sum-of-square differences (SSD) or normalised cross-correlation (NCC) because it takes into account the fact that the noise in an ultrasound image is multiplicative Rayleigh noise, and that displayed ultrasound images are log-compressed. However it assumes that the noise distribution in both of the blocks is the same and this assumption is not correct for ultrasound images. The attenuation of the ultrasound waves introduces inhomogeneities in the image of homogeneous tissue [15]. The time gain and the lateral gain compensations (compensating respectively for the effects that deeper tissue appears dimmer and for intensity variations across the beam) which are tissue independent and generally constant for a given location during the acquisition, do not compensate fully for the attenuation. Further, because of the large velocity dynamic of the myocardial wall in an echocardiographic sequence, the time gain and the lateral gain compensations will be different for the same tissue for different frames in the sequence. Thus in this work an intensity normalisation is conducted before calculation of the CD_2 similarity measure. This is achieved by making sure that the two blocks of data have at least the same mean and variance. In more detail, the original intensity values $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{y}}_i$ above are transformed into new values of $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{y}}_i$ by subtracting the mean and dividing by the standard deviation (square root of the variance) of the intensity values in the block. We call this the modified similarity CD_{2bis} .

3 Calculation of optical flow estimates (Singh approach)

Having defined the similarity measure, we now address how it is used in a block-matching framework for the estimation of the velocity field. In the previous section, we used vectorial notation, however, in this section detailed notation is needed as we have to distinguish between the two directions x and y .

¹ Note eq. 10 was given in [4] using the data before log compression (i.e. x_{ij} and y_{ij}).

Due to the well-known aperture problem, velocity estimation using eq. 1 can recover only the component of velocity normal to intensity edges. The full velocity estimate can be calculated by propagating information from regions or “interesting points” that do not suffer from the aperture problem, such as corners. Moreover, the displacement vector \mathbf{v} is discretised since the search region is discrete, and hence the accuracy of the velocity field is limited by this discretisation. To obtain more reliable estimates a smoothness constraint on the velocity field with an appropriate confidence measure must be incorporated. In Singh’s approach both conservation information (a data constraint) and neighbourhood information (a smoothness constraint) are combined in a probabilistic framework based on estimates of their associated errors [13].

In the first step (conservation information), a square window W_c having a side length of $2n + 1$ is formed about a central pixel (x, y) in the first frame. A square search window W_s of side length $2N + 1$ is defined in the second frame around the position of the corresponding central pixel (x, y) in the second frame. The size of W_s depends on the assumed maximal displacement between two successive frames.

The intensities of the block W_c of pixels in frame 1 are then compared with all possible positions of the block in the search window W_s using a similarity measure. This gives a value of $E_c(u, v)$ for each candidate displacement (u, v) . Thus in a first step based on conservation information the similarity values E_c are used in a probability mass function to calculate a response R_c whose value at each position in the search window represents the likelihood of the corresponding displacement. Singh used the SSD as a similarity measure and the following function as a probability density function:

$$R_c(u, v) = \frac{1}{Z} \exp\{-kE_c(u, v)\} \quad -N \leq u, v \leq N, \quad (12)$$

where Z is defined such that all probabilities sum to one and the parameter k is chosen at each position such that the maximum response in W_s is close to unity (0.95 before normalisation) for computational reasons. Singh then defines a velocity estimate as being the mean of the probability mass function:

$$\mathbf{v}_c = \begin{pmatrix} u_c \\ v_c \end{pmatrix} = \begin{pmatrix} \sum_u \sum_v R_c(u, v)u \\ \sum_u \sum_v R_c(u, v)v \end{pmatrix}, \quad (13)$$

and its associated error (called the conservation error) by:

$$(\mathbf{v} - \mathbf{v}_c)^T \mathbf{S}_c^{-1} (\mathbf{v} - \mathbf{v}_c), \quad (14)$$

where, \mathbf{S}_c is the covariance matrix given by:

$$\mathbf{S}_c = \begin{pmatrix} \sum_{u,v} R_c(u, v)(u - u_c)^2 & \sum_{u,v} R_c(u, v)(u - u_c)(v - v_c) \\ \sum_{u,v} R_c(u, v)(u - u_c)(v - v_c) & \sum_{u,v} R_c(u, v)(v - v_c)^2 \end{pmatrix}. \quad (15)$$

Another velocity estimate may be obtained by the use of neighbourhood information. In other words, the velocity at each pixel is unlikely to be completely

6 Djamel Boukerroui et al.

independent of the velocity of its neighbours. Thus, assuming that the velocity of each pixel in a small neighbourhood W_p has been estimated, the velocity for each pixel can be refined by using the velocity of its neighbouring pixels. Clearly it is more likely that the velocities of closer neighbours are more relevant than pixels which are further away. Therefore weights are assigned to velocities calculated for the neighbouring pixels, and the weights drop with increasing distance from the central pixel (a 2-D Gaussian mask in the window W_p of size $(2w+1)(2w+1)$ is used). These weights can be interpreted as a probability mass function $R_n(u_i, v_i)$ where i is an index for pixels in W_p . In the same way as for the conservation information a second velocity estimates is obtained $\mathbf{v}_n = (u_n, v_n)^T$ with its associated covariance matrix \mathbf{S}_n .

The sum of conservation (data term) and neighbourhood (regularisation term) errors represents the total squared error of the velocity estimate:

$$\epsilon^2(\mathbf{v}) = (\mathbf{v} - \mathbf{v}_c)^T \mathbf{S}_c^{-1} (\mathbf{v} - \mathbf{v}_c) + (\mathbf{v} - \mathbf{v}_n)^T \mathbf{S}_n^{-1} (\mathbf{v} - \mathbf{v}_n) . \quad (16)$$

The optimal velocity is that which minimises this error and can be obtained by setting the gradient of the error with respect to \mathbf{v} to zero giving:

$$\hat{\mathbf{v}} = [\mathbf{S}_c^{-1} + \mathbf{S}_n^{-1}]^{-1} [\mathbf{S}_c^{-1} \mathbf{v}_c + \mathbf{S}_n^{-1} \mathbf{v}_n] . \quad (17)$$

Because \mathbf{v}_n and \mathbf{S}_n are derived on the assumption that the velocity of each pixel of the neighbourhood is known in advance, in practice equation (17) is solved in an iterative process (Gauss-Seidel relaxation) with the initial values of the velocity at each pixel being taken from the conservation information alone:

$$\begin{cases} \hat{\mathbf{v}}^0 &= \mathbf{v}_c \\ \hat{\mathbf{v}}^{m+1} &= [\mathbf{S}_c^{-1} + (\mathbf{S}_n^m)^{-1}]^{-1} [\mathbf{S}_c^{-1} \mathbf{v}_c + (\mathbf{S}_n^m)^{-1} \hat{\mathbf{v}}_n^m] \end{cases} . \quad (18)$$

Inspecting the energy function to be minimised (eq. 16) we notice that the parameter k plays an important role as the covariance matrix \mathbf{S}_c is highly dependent on the value of k . Indeed, k controls the contribution of the conservation information in the total energy and thus the amount to which \mathbf{v}_c is regularised using the neighbourhood information. In [13], the value of k is estimated at every pixel position implying that the regularisation is not uniform over the whole velocity field. A second weak point of the Singh approach is in taking the expectation of the probability as the expected velocity introduces errors where the probability mass function is non-modal. A multiple modal pdf does occur mainly because of the aperture problem. However, a low signal-to-noise ratio is a second source. To avoid these two limitations, we define the probability mass function as follow:

$$R_c(u, v) = \frac{1}{Z} \exp \left\{ k \frac{E_c(u, v) - m}{(2n+1)^2} \right\} \quad -N \leq u, v \leq N , \quad (19)$$

where E_c is the CD_{2bis} similarity measure and m is its maximum in the search windows W_s . Notice that the maximum of R_c , before normalisation, has a value

of one by construction; this avoids numerical instability problems. In our formulation, the parameter k has a constant value leading to a more uniform regularisation of the velocity field. The similarity measure is normalised by the size of the correlation windows W_c . Hence the parameter k is independent of the size of W_c so that the same value of k could be used for different window sizes. This is particularly very effective when using a coarse-to-fine estimation strategy.

One way to get more reliable estimates of the expected velocity when the pdf is not mono-modal is to bias the estimation towards the predominant mode. In this work we take as an estimate of the velocity, the mean of a thresholded version of the probability mass function. In other words:

$$u_c^h = \frac{\sum_u \sum_v R_c^h(u, v) u}{\sum_u \sum_v R_c^h(u, v)} , \quad v_c^h = \frac{\sum_u \sum_v R_c^h(u, v) v}{\sum_u \sum_v R_c^h(u, v)} , \quad (20)$$

where,

$$R_c^h(u, v) = \begin{cases} R_c(u, v) & \text{if } R_c(u, v) \geq \alpha \\ 0 & \text{otherwise} \end{cases} ; \quad (21)$$

and the threshold α is defined as follow:

$$\alpha = \widehat{m} - h(\widehat{m} - \widetilde{m}) \quad \text{with } h \in [0, 1] , \quad (22)$$

where \widehat{m} and \widetilde{m} are the maximum and the minimum of the probability mass function R_c respectively.

Equation (20) is equivalent to equation (13) for $h = 1$ (i.e. the velocity is the mean), and is equivalent to taking as the velocity, the argument of the maximum of the probability mass function for $h = 0$. One can optimise the value of h , given a ground true data for a given type of images, to have the best estimation with subpixel precision. This has the advantage to overcome the lack of robustness to noise and precision of the maximum approach and solve the problem of the multimodal pdfs for the mean approach as the new estimates will be biased towards the predominant mode. Note that the covariance matrix is calculated using the whole probability around the new estimates (which is different from the mean using the whole pdf). Therefore, the error in the velocity estimates does take into account the presence of a second mode. And, the second mode information is also taken into account in the regularisation framework.

4 Parameters optimisation and implementation details

In this section we briefly describe the optimisation scheme for the estimation of k and h and give the reader the important implementation details.

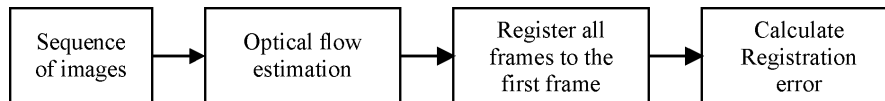


Fig. 1. Block diagram of the proposed parameters optimisation scheme.

8 Djamal Boukerroui et al.

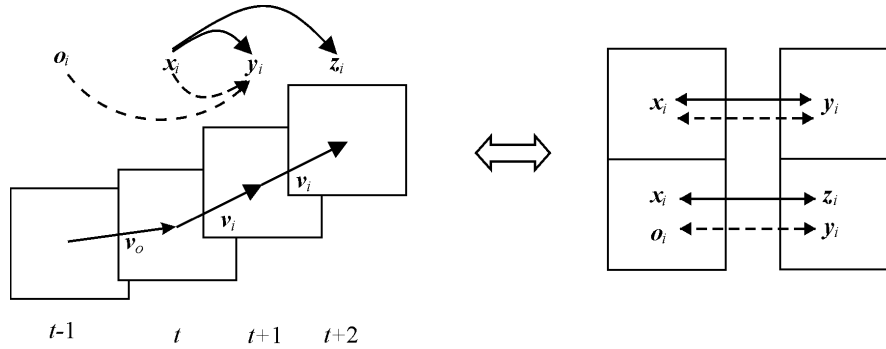


Fig. 2. Illustration of the 3 frames scheme for the estimation of the velocity between frame x at time t and frame y at time $t + 1$. The constant velocity approach uses the frame at $t + 2$ (continuous arrows); the proposed approach uses the preceding frame at time $t - 1$ (dotted arrows).

Parameters optimisation: k and h may be optimised using a simulated sequence where the true velocity field is known. As we noted, k and h are highly dependent both on the signal-to-noise ratio and on the velocity field to be estimated (i.e. the amount of the smoothness of the field). Therefore, we choose an optimisation scheme using the original data. Figure 1 illustrates schematically how the values of $k \in]0 \infty[$ and $h \in [0 1]$ are optimised.

A velocity field is estimated using initial values of k and h , and then all the subsequent frames are registered to the first frame using the calculated cumulative optical flow field. Finally, a registration error $\mathcal{E}(k, h)$ is calculated using a similarity measure. The SSD similarity is used in our experiments and we defined the optimal parameter set as:

$$(\hat{k}, \hat{h}) = \arg \min_{(k, h)} \mathcal{E}(k, h) . \quad (23)$$

The Powell multidimensional minimisation algorithm was used to solve the above optimisation problem. We found the method relatively robust to the choice of initial values, for example, we found initial values of $h = 0.5$ and $k = 0.5$ suitable for an ultrasound imaging sequence.

Implementation details: A common problem in optical flow estimation using matching techniques is the multi-modal response (i.e due to the aperture problem for example or a mismatching specially when the size of the search windows is large). A common way to reduce its effect is to make the assumption of a local stationary flow field (usually 3 frames). This assumption is relatively true for high frame rate data and a smooth velocity field over time. Unfortunately, in the case of contrast echocardiography, tissue Doppler and real-time 3D-imaging low frame rates are typical (20-30 Hz). We suggest an alternative approach to tackle this problem without making any assumption on the velocity field, but by assuming that the observed moving tissue conserves its statistical behaviour through time (at least for 3 to 4 consecutive frames).

Suppose $\mathbf{o}, \mathbf{x}, \mathbf{y}$ and \mathbf{z} are four consecutive frames respectively at times $t - 1, t, t + 1$ and $t + 2$. Figure 2, illustrates the blocks being compared for the two approaches (the constant velocity and the proposed one). Our approach makes use of the calculated velocities between the preceding frame \mathbf{o} and the current frame \mathbf{x} . Given a block \mathbf{x}_i in frame \mathbf{x} at time t , which is compared to blocks \mathbf{y}_i in the search window W_s in frame \mathbf{y} at the time $t + 1$; using the previous estimated velocity, we track back the position of \mathbf{x}_i in the preceding frame at time $t - 1$ and we denote the corresponding block by \mathbf{o}_i . Hence, theoretically \mathbf{x}_i and \mathbf{o}_i could be seen as two independent observations of the same tissue. Thus in the new approach, intensities of each candidate block \mathbf{y}_i in the search window W_s are compared with the intensities of the block \mathbf{x}_i centred at (x, y) in the frame \mathbf{x} at time t , and also with the corresponding block \mathbf{o}_i centred at the calculated position $(x - u_o, y - v_o)$ in frame \mathbf{o} at time $t - 1$, where $\mathbf{v}_o = (u_o, v_o)^T$ is the displacement field of \mathbf{o}_i to \mathbf{x}_i .

Finally, a coarse-to-fine strategy is used to reduce the computational load of the algorithm when the expected velocity range is large. A multiresolution implementation is used, as suggested by Singh [13] to which the reader is referred for more details.

5 Results

The registration error $\mathcal{E}(k, h)$ surface of one experiment conducted on the ultrasound breast data is shown in Fig.3. Three important observations can be made:

1. For $h = 0$, the velocity estimation is equivalent to taking the argument of the maximum of the pdf. Hence, theoretically, the parameter k does not have any influence on the result. This can easily be observed for this experiment, and it corresponds to the maximum error. In this case, the optical flow is quantified by the pixel resolution of the image, and hence the error on the image velocity is of the order of the pixel resolution. Furthermore, this

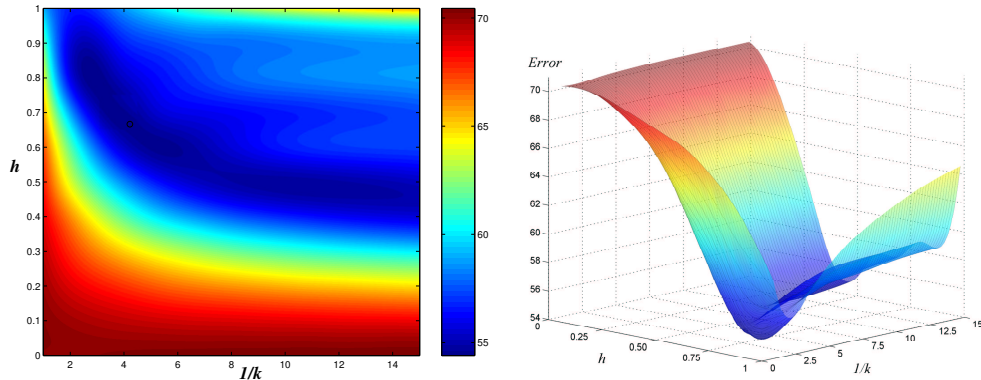


Fig. 3. The registration error $\mathcal{E}(k, h)$ for a breast ultrasound sequence.

10 Djamel Boukerroui et al.

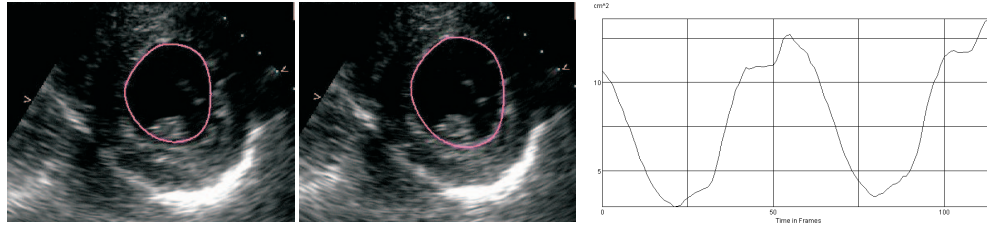


Fig. 4. Endocardial wall tracking using velocity estimation. First frame (left); last frame (middle); Area plot of the left ventricle function of time (right).

approach is not robust against noise. This explains the high error on the velocity estimation.

2. For $h = 1$ (as in Singh approach), the velocity estimation is equivalent to taking the mean of the probability. The results are better than for $h = 0$, but does not correspond to the optimal value. This result can be explained by the fact that taking the mean of the probability as an estimates of the velocity is not very precise and may conduct to biased estimation if the pdf is not mono-modal or non-well-peaked pdfs. Observe as well the expected functional dependence between the two parameters (h and k). Therefore, the search for the optimal values of h and k must be done in the 2D space.
3. Inspecting the results we notice that for $h = 1$ (as in Singh), the optimisation of k makes an improvement of about 15% relative to the worst value for k . The optimisation of h for different values of k makes an enhancement between 20% to 35%.

In the above experiment the optimal values are $h = 0.660$ and $k = 0.237$. Using these values, an enhancement of approximately 35% is achieved in comparison to result obtained using the worst parameters set.

Figure 4 shows a tracking example of cardiac boundary pixels on a short axis echocardiographic sequence. In this example, the velocity estimation is done only at contour points (i.e we mean both steps: the similarity calculation as well as the regularisation) which reduces the computational burden. The area plot of two cardiac cycles (over 100 frames) demonstrates the subpixel accuracy of the method. Indeed as the errors on the velocity estimation are propagated from frame to frame, a poor estimation accuracy will result in a large errors after one cardiac cycle.

Figure 5 shows a second example of motion tracking on a free hand ultrasound breast data. The figure shows four frames at regular interval of the sequence (about 300 frames). Tissue motion estimation on this type of data is very difficult. Notice for example that at the end of the sequence, there is no signal at the right hand side of the images. As in the previous example, the velocity estimation is done only at contour points. We used one frame out of five frames acquired to demonstrate the robustness of the method in the case of low frame rate data. Without using any high level processing (global motion model, more constrained contour model) the proposed block matching provides satisfactory tracking results. The result are more convincing when visualised as a movie.

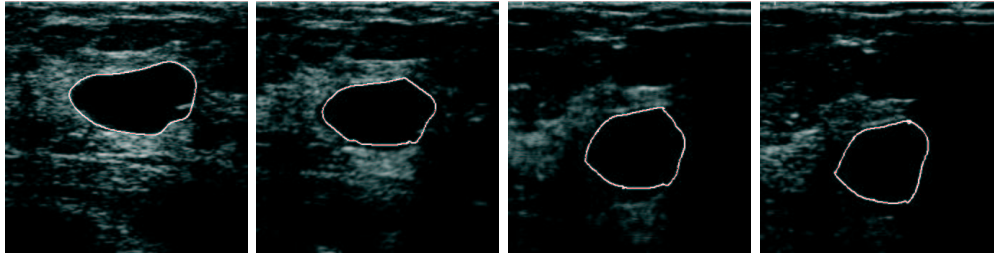


Fig. 5. Breast mass tracking using velocity estimation. The images are at different times in the sequence. Here only one frame every 5 frames acquired is used. Notice the shadow on the right hand side of the last two images.

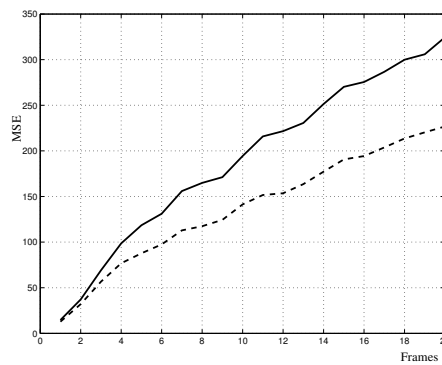


Fig. 6. Evolution of the Mean Square Error between the first frame (0) and the registered one to the first frame. Here a dense velocity field is calculated. Results obtained using: the NCC similarity (solid curve); the CD_{2bis} similarity (dotted curve).

The last experiment presents a comparison of two similarity measures (the Normalised Cross Correlation and the CD_{2bis}) used within this framework for the estimation of a dense velocity field. The breast data are used in this experiment. Figure 6 shows the evolution of Mean Square tracking Error over the frames. The MSE is calculated between the registered frames and the first frame using the cumulative velocity field. The results demonstrate the superiority of the CD_{2bis} similarity measure.

All experiments are obtained with a side length of 21 pixels (≈ 7 mm) for the correlation window W_c and 19 pixels for the search window W_s . With the above windows sizes and for two levels for the coarse-to fine strategy, the computation time is about 3 frames per second per contour on a Pentium 1.8 GHz.

6 Conclusion

In this paper a novel combination of two existing techniques (namely a similarity measure proposed by Cohen and Dinstein and the Singh block matching approach) is proposed. Improvements on both techniques are made leading to more effective block-matching algorithm in ultrasound sequences. A global optimisation scheme for the parameters estimation of the algorithm is proposed.

12 Djamel Boukerroui et al.

It is demonstrated that this optimisation improves the result of about 35% in comparison to the worst case. In this paper, the performance of the proposed block matching approach is demonstrated on B-mode ultrasound images from two application areas. We obtained reliable boundary tracking without use of any high level constraints on the motion field.

Acknowledgements: We are grateful to Dr. M. M. Parada and Dr. J. Declerck, from Mirada Solutions Ltd, for providing software used in part of this work. This work was funded by the EC project ADEQUATE (IST: 1999-10837) and the UK GPSRC funded IRC in Medical Image Analysis and Signals (MIAS).

References

1. Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. *Int. Journal of Computer Vision*, **12**(1994) 43–77.
2. Bernard, C.: Discrete wavelet analysis for fast optic flow computation. Internal Report RI415, Centre de Mathématiques Appliquées, Ecole Polytechnique, France, (1999).
3. Baxley, W.A., Reeves T.J.: Abnormal regional myocardial performance in coronary artery disease. *Prog Cardiovasc Dis* **13**(1971) 405–421.
4. Cohen, B., Dinstein, I.: New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition* **35** (2002) 455–463.
5. Han, L., Burcher, M. Noble, J.A.: Non-invasive measurement of biomechanical properties of in vivo soft tissues. In *Proc. MICCAI*, (2002) 208–215.
6. Kerber R.E., Abboud F.M.: Echocardiographic detection of regional myocardial infraction. *Circulation* **47** (1973) 997–1005.
7. Lai, S.A., Vemuri, B.C.: Robust and efficient computation of optical flow. *International Journal of Computer Vision* **29** 2 (1998) 87–105.
8. Ledesma-Carbayo, M.J., Kybic, J., Desco, M., Santos, A., Unser, M.: Cardiac motion analysis from ultrasound sequences using non-rigid registration. In *Proc. Int.Conf. Medical Image Computing and Computer-Assisted Intervention*, (2001) 889–896.
9. Ledesma-Carbayo, M.J., Kybic, J., Sühling, M., Hunziker, P., Desco, M., Santos, A., Unser, M.: Cardiac ultrasound motion detection by elastic registration exploiting temporal coherence. In *Proc. IEEE Int. Symposium on Biomedical Imaging: Washington DC, USA*, (2002) 585–588.
10. Mikic, I. Krucinski, S., Thomas, J. D.: Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates. *IEEE Trans. Med. Imaging* **17** 2 (1998) 274–284.
11. Mulet-Parada, M., Noble, J. A.: 2D+T acoustic boundary detection in echocardiography. *Medical Image Analysis* **4** (2000) 21–30.
12. Mulet-Parada, M.: Intensity independent feature extraction and tracking in echocardiographic sequences. Phd thesis, MVL, Dept. of Eng. Science. Oxford: Oxford University, (2000).
13. Singh, A.: Image-flow computation: An estimation-theoretic framework and a unified perspective. *CVGIP: Image understanding* **65** 2 (1992) 152–177.
14. Weber, J. Malik, J.: Robust computation of optical flow in a multi-scale differential framework *International Journal of Computer Vision* **14** 1 (1995) 67–81.
15. Hughes, D. I., Duck, F. A.: Automatic attenuation compensation for ultrasonic imaging. *Ultrasound in Medicine & Biology*, **23** (1997) 651–664.

Identification automatique des régions d'intérêts sur des images tomographiques rénales

Identification automatique des régions d'intérêts sur des images tomographiques rénales

Automatic identification of regions of interest on renal tomographic images

Djamal BOUKERROUI¹, Wala TOUHAMI², Jean Pierre COCQUEREZ¹

¹Université de Technologie de Compiègne, CNRS UMR 6599 Heudiasyc
BP 20529 - 60205 Compiègne Cedex, France

²École Nationale d'Ingénieurs de Tunis
BP 37, Le Belvedere 1002 Tunis

djamal.boukerroui@hds.utc.fr, wala.touhami@gmail.com, jean-pierre.cocquerez@hds.utc.fr

Résumé — Nous proposons, dans le présent papier, une approche originale dans un cadre statistique pour l'identification automatique des reins (sains et pathologiques) sur des images tomographiques bidimensionnelles (CT). Notre approche est constituée de deux phases: une phase de localisation suivie d'une phase de délimitation. La phase de localisation est guidée, d'une part, par un modèle *a priori* spatial et d'autre part, par un modèle *a priori* sur les niveaux de gris, statistiquement appris. La seconde phase consiste à utiliser les résultats de la localisation afin de délimiter la région du rein en utilisant un ensemble de règles. Cette approche est testée sur des images cliniquement acquises et des résultats satisfaisants sont obtenus.

Détection automatique, approche statistique, modèles *a priori*, kyste de rein, CT.

Abstract — We propose, in this paper, an original approach in a statistical framework, for fully automatic delineation of kidneys (healthy and pathological) in 2D CT images. Our approach has two main steps: a localisation step followed by a delineation step. The localisation step is guided by a statistically learned *prior* spatial model in one hand and a grey level *prior* model in a second hand. The second step, utilizes the localisation results in order to precisely delineate the kidney's regions using a set of learned IF-THEN rules. The proposed approach is tested on clinically acquired images and promising results are obtained.

Automatic detection, statistical approach, *prior* models, kidney cysts, CT.

1 Introduction

Le développement et la prolifération rapide des technologies de l'imagerie médicale révolutionne la médecine. L'imagerie médicale permet aux scientifiques et aux médecins de recueillir des informations en scrutant d'une façon non invasive dans le corps humain. Le rôle de l'imagerie médicale est passé au delà de la visualisation simple et de l'inspection des structures anatomiques. Elle est devenue un outil pour la planification et la simulation chirurgicales, pour le dépistage des kystes et des tumeurs, pour le suivi de la progression des maladies, etc. Pour ces raisons, la communauté d'analyse d'images médicales s'est préoccupée du problème fastidieux de l'extraction et de l'identification des structures anatomiques. En effet, les processus de segmentation des images médicales et d'identification des structures anatomiques sont des étapes primordiales. Les images scanner de l'abdomen sont particulièrement

difficiles à traiter automatiquement pour de nombreuses et diverses raisons. Sans soucis d'exhaustivité, nous citons les deux plus importantes. Malgré les récents progrès en Tomodensitométrie, les images cliniquement acquises sont caractérisées par la présence d'artefacts du processus de reconstruction, la présence d'artefacts de volume partiel et par un faible contraste entre les frontières de certains organes de densité électronique similaire. Ces caractéristiques écartent la possibilité d'utiliser les techniques standards, telles que la détection de contours ou de classification [15, 31], qui agissent d'une façon aveugle sur les niveaux de gris. La seconde difficulté majeure est anatomique. Elle est due au caractère mou et complexe des différents organes de l'abdomen, introduisant une grande variabilité des structures selon les images et les patients. Cette variabilité augmente sensiblement la difficulté de l'identification automatique des organes et les méthodes traditionnelles de reconnaissance des formes échouent. De

plus, cette variabilité anatomique est augmentée par la présence de la pathologie ou par l'injection de produits de contraste. Cette problématique invite donc à une réflexion profonde sur les méthodes de traitements à utiliser. De nombreuses méthodes ont été proposées. Toutes reposent sur une utilisation plus efficace de l'*a priori* sur le problème à résoudre. Dans ce qui suit, nous présentons brièvement un état de l'art des travaux les plus importants qui se rapprochent de notre problématique.

1.1 État de l'art

A notre connaissance, Karssemeijer et al. sont les premiers à avoir proposé une approche exploitant des connaissances anatomiques pour extraire les organes sur des images CT de l'abdomen [20]. Une caractéristique importante de la méthode est l'utilisation de zones de recherche pour guider le processus de segmentation d'une structure anatomique donnée. La méthode a été appliquée pour la reconnaissance de la rate avec des résultats encourageants qui suggèrent l'efficacité d'une telle approche.

Kobashi et al. [21] ont étudié un système basé sur les connaissances anatomiques pour extraire les organes sur des images CT de l'abdomen. Ils proposent une approche algorithmique composée de deux modules : un module connaissances et un module moteur. Le premier contient les propriétés des organes : leur densité, position, voisinage, etc. Le second gère les méthodes de traitement d'images et d'extraction d'attributs. Le processus d'extraction des organes est incrémental, c'est à dire que chaque organe est extrait de toutes les coupes avant de procéder à l'extraction d'un autre organe. Pour cela, les auteurs adoptent une approche reposant sur un seuillage dynamique contrôlé en retour par les propriétés des régions résultantes (aire, position, largeur, etc.).

Archip et al. [1] proposent une approche semi-automatique, utilisant une base de connaissances *a priori*, pour la segmentation de la moelle épinière sur des séquences d'images CT. Le principe de leur système se base sur la construction d'une carte anatomique des structures (Anatomical Structures Map). Cette carte est constituée d'un ensemble de propriétés caractérisant les organes et les tissus sains (e.g. position relative, forme, plage de densité). L'arrangement spatial des différents organes est également représenté par un réseau sémantique exploitant des relations simples (e.g. Est-il à gauche, est-il à l'intérieur, etc.). Ensuite, cette carte est introduite dans une architecture, appelée "plan solver", qui oriente le système vers la méthode de segmentation adéquate en utilisant les propriétés définies dans la carte anatomique (seuillage pour la détection des os ou croissance de régions pour le canal vertébral et la moelle épinière).

L'équipe de Kaneko a une préférence pour l'utilisation des outils proposés par la théorie de la morphologie mathématique pour résoudre le problème de segmentation

et de reconnaissance des organes abdominaux. Un nouvel outil, le chapeau haut de forme différentiel est proposé et utilisé pour la segmentation automatique des différentes structures abdominales. Une étape de séparation et de reconnaissance d'organes est effectuée basée essentiellement sur des statistiques *a priori* de position et de taille. La méthode est testée sur 4 séquences d'images [19]. Cette approche nous semble intéressante mais nécessite de plus amples expérimentations.

Lee et al. [23] proposent un système d'identification de différents organes abdominaux, présents sur des images tomographiques. Leur approche est basée sur des règles floues. En premier lieu, les niveaux de gris et l'information contextuelle de voisinage sont utilisés pour classer les pixels de l'image en utilisant un réseau de neurones. Ensuite, un ensemble de descripteurs flous de forme et de position (position relative, distance relative, compacité, etc.) est calculé sur les régions segmentées afin de reconnaître les différents organes.

Les relations spatiales des différentes structures abdominales sont également utilisées dans un algorithme hiérarchique de segmentation dans [6]. Dans cette approche, l'extraction d'un organe repose sur des informations issues d'autres informations plus simples à obtenir. En d'autres termes, l'algorithme exploite les organes déjà détectés afin de restreindre l'espace de recherche des structures restantes. Ainsi la peau, les structures osseuses, les poumons, les reins puis le foie sont détectés automatiquement dans l'ordre.

Très récemment, la même équipe a présenté l'utilisation du formalisme flou pour la modélisation des relations spatiales, dans un algorithme automatique de segmentation du coeur, comme organe, sur des images CT [27]. L'originalité de ce travail réside dans l'emploi de ce formalisme pour la modélisation des connaissances anatomiques du coeur, en particulier sa position par rapport aux deux poumons. Cette connaissance est d'abord utilisée pour définir une région d'intérêt. Ensuite, un modèle déformable intégrant ces connaissances anatomiques, est initialisé dans la région d'intérêt pour segmenter le coeur. Le succès de cette méthode repose essentiellement sur la possibilité d'identifier les poumons sans difficulté et de pouvoir, par conséquent les utiliser comme une initialisation des modèles déformables. Par ailleurs, la stabilité relative de la position du coeur entre les poumons a permis l'élaboration d'un modèle de relations spatiales fiable pour la segmentation.

Park et al. [30] proposent une approche de segmentation d'images tomographiques 3D dans laquelle l'*a priori* est exprimé sous forme d'un atlas probabiliste des organes abdominaux. La segmentation est effectuée dans un cadre bayésien avec un *a priori* markovien exploitant l'information de l'atlas. L'association d'un atlas est largement répandue pour résoudre la segmentation du cerveau, mais son utilisation devient très problématique sur les régions abdominales car la construction de l'atlas nécessite

la résolution d'un problème de recalage très difficile (voir e.g. [7, 42]). Dans ce travail, les auteurs ont proposé d'effectuer un recalage pour chaque organe séparément, ce qui nécessite l'intervention de l'utilisateur pour la détection de l'organe. L'atlas probabiliste, couplé avec une méthode de segmentation du type ensembles de niveaux, améliore les résultats [34].

Une autre façon de résoudre des problèmes de segmentation difficiles est de faire appel aux Modèles Actifs de Formes (MAF), initialement introduit dans [8, 9]. L'idée est de rechercher la solution dans un espace plus restreint en utilisant un modèle de forme paramétrique, obtenu par l'Analyse en Composantes Principales (ACP) d'un ensemble d'exemples représentatifs de la forme recherchée. Le problème de mise en correspondance des points, qui est un des inconvénients majeurs des MAF, a été solutionné –plus précisément évité– par le modèle implicite de forme de Leventon et al. [24]¹. Dans ce dernier, l'ACP est appliquée sur des représentations implicites des formes de l'objet, i.e en utilisant la fonction distance signée. Ainsi, ils s'adaptent mieux aux récentes méthodes de segmentation variationnelles (voir [11] pour un état de l'art sur les différentes méthodes du domaine). L'application de ces modèles a permis de résoudre de nombreux problèmes de segmentation en imagerie médicale (e.g. [16, 24, 35, 40, 41, 43]). Cependant les approches qui imposent des *a priori* statistiques de formes, souffrent de certaines limitations, notamment l'hypothèse gaussienne de la distribution des formes de l'ensemble d'apprentissage. Quelques travaux ont été faits pour pallier cet inconvénient majeur [10].

Très récemment, une autre solution robuste au problème de mise en correspondance des points dans la création du modèle de forme, a été utilisée avec succès pour la segmentation 3D des reins sur des images CT. Une base de données de 41 acquisitions est exploitée [35]. L'information de forme, modélisée d'une façon plus classique, i.e via un vecteur de caractéristiques locales, apparaît également dans [39], pour la segmentation des reins sur des images scanner 3D. Un modèle déformable paramétrique est utilisé. Cette approche a la particularité d'être automatique. Une méthode d'initialisation automatique du modèle déformable est proposée. Elle consiste principalement en deux étapes : la première définit une position *a priori* des régions des reins par rapport à la région abdominale ; la seconde recherche une position plus précise du rein en utilisant un motif (template). Lin et al. [25] proposent une variante de l'étape de localisation de Tsagaan et al. [39] pour localiser un germe dans le rein, puis ils utilisent un simple algorithme de croissance de région pour le segmenter. Les statistiques des résultats sont encourageantes, cependant, dans toutes les images de test les reins

sont fortement contrastés.

Nous avons résumé ci-dessus quelques travaux développés récemment pour l'identification ou la segmentation avec une modélisation de l'information *a priori*. Notons que deux types d'information sont essentiellement utilisées. Une information de position de l'objet d'intérêt et une information sur la forme et/ou l'apparence de l'objet. Un autre point partagé par les approches présentées est leur application commune aux problèmes de segmentation où la variabilité de la région recherchée est limitée, i.e. des organes sains ou dont la pathologie n'introduit pas une grande variabilité de forme ou d'apparence. La nécessité de l'intervention de l'utilisateur, dans certaines d'entre elles, est également problématique pour notre application. Enfin, les quelques approches automatiques existantes, travaillent sur des données tridimensionnelles.

Dans ce travail, nous nous intéressons à identifier automatiquement les régions des reins sains ou pathologiques sur des images CT bidimensionnelles, avec ou sans produit de contraste. Nous entendons par pathologique, des reins avec un ou plusieurs kystes. Comme nous allons le voir sur les exemples présentés, les kystes introduisent une grande variabilité de forme et d'apparence, qui nous a contraint à abandonner les solutions présentées ci-dessus. Le fait de travailler sur des images 2D, augmente la variabilité de position et de forme du rein. Enfin, rappelons que notre but est seulement de détecter au mieux tout le rein et les kystes. Ainsi, nous tolérerons un rein grossièrement détecté mais pas partiellement détecté.

2 Architecture globale de l'approche proposée

La méthode proposée pour l'identification automatique des reins s'inscrit dans un cadre statistique. L'observation des images CT de l'abdomen de notre base de données, ainsi que les différentes discussions échangées avec les radiologues nous ont permis de dégager différents descripteurs pertinents caractérisant les reins. En effet, les reins se situent généralement de part et d'autre de la colonne vertébrale et souvent dans la partie inférieure de l'abdomen, sauf dans certains cas où le rein est greffé. Par ailleurs, ils ont en moyenne une densité de 30 Unités Hounsfield (UH) sur les tissus sains, sans kyste ni tumeur. Ces propriétés suggèrent la possibilité de guider leur identification en utilisant un modèle *a priori* sur la position spatiale et un modèle *a priori* sur les niveaux de gris. La question qui se pose ensuite est de savoir utiliser les modèles construits pour détecter les reins. Pour y répondre, nous proposons une approche originale et automatique. Elle est composée de deux phases : Une première phase de localisation des reins et une deuxième phase de détection.

– **Phase de localisation des reins** : L'utilisation d'un modèle statistique de la position spatiale et des

¹D'autres solutions existent bien sûr. Nous pensons par exemple aux travaux de Cootes et Taylor qui proposent l'application d'algorithmes de sélection de modèles (du type MDL). Voir également [35] pour une proposition plus récente.

niveaux de gris des régions d'intérêt est la clé de succès de cette phase.

- **Phase de détection des reins :** Ici, tout le rein est détecté par une méthode de type agrégation de régions en partant de celles détectées dans la première phase.

3 Phase de localisation des reins

Elle est basée sur l'estimation de modèles statistiques *a priori* de la position spatiale et des niveaux de gris, qui seront utilisés pour conduire le processus de localisation. Le schéma de la figure 1 présente cette phase composée essentiellement de deux étapes :

- Une étape de construction des modèles statistiques *a priori* : elle permet de modéliser les deux caractéristiques discriminantes des reins, à savoir la position spatiale et la distribution des niveaux de gris. Un ensemble d'apprentissage d'images CT de l'abdomen permet d'estimer un premier modèle *a priori* spatial et un second modèle *a priori* sur les niveaux de gris des reins.
- Une étape de localisation : le problème de la localisation prend la forme d'un problème d'optimisation d'un critère prenant en compte les modèles *a priori* déjà construits pour localiser deux fenêtres sur les régions des reins.

Nous décrivons en détail ces deux étapes dans les paragraphes suivants.

3.1 Modélisation de l'information *a priori*

La modélisation de l'*a priori* constitue la clé de notre approche d'identification automatique des reins. Nous souhaitons modéliser la variation de la forme ainsi que la variation des niveaux de gris des reins, à partir d'un ensemble d'apprentissage constitué de N images abdominales noté $A = \{A_i, i = 1 \dots N\}$. Comme nous nous intéressons aux reins gauche et droit, la notation (g, d) est adoptée pour les désigner.

Soient Ω le support spatial d'une image et \mathcal{L} l'ensemble des niveaux de gris. Nous définissons un champ aléatoire $Y = \{y_s, s \in \Omega, y_s \in \mathcal{L}\}$, décrivant les observations et un

champ aléatoire binaire $X = \{x_s, s \in \Omega\}$ tel que :

$$x_s = \begin{cases} 1 & \text{si le pixel } s \text{ appartient au rein;} \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, l'image A_i est la réalisation y^i du champ aléatoire Y . Une réalisation x^i du champ X peut être construite en segmentant manuellement les régions des reins en utilisant les données y^i .

3.1.1 Modèle *a priori* spatial

Les régions des reins segmentées manuellement sont utilisées pour modéliser empiriquement l'*a priori spatial* en calculant le champ moyen en chaque site s :

$$p(s) = \frac{1}{N} \sum_{i=1}^N \frac{x_s^i}{n_i}, \quad (1)$$

où n_i est le nombre de pixels de l'image A_i tel que $x_s^i = 1$. Ainsi, nous pouvons construire un modèle *a priori* pour les reins gauche et droit, à partir des images de l'ensemble d'apprentissage.

3.1.2 Modèle *a priori* des niveaux de gris :

Soit ω une variable aléatoire, qui prend sa valeur dans \mathcal{L} . De façon similaire au modèle spatial, nous pouvons construire empiriquement un modèle *a priori* de la distribution des niveaux de gris à partir des images des reins segmentées manuellement. Notons que cet *a priori*, désigné par $p_\omega(\cdot)$, est indépendant de la position du pixel dans l'image et il est donné par :

$$p_\omega(l) = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{n_i} \sum_{\{s/x_s^i=1\}} \delta_l(y_s^i) \right\}, \quad (2)$$

ici, y_s^i est la valeur du niveau de gris du pixel s de la $i^{\text{ème}}$ image et $\delta_l(\cdot)$ est la fonction delta en l , définie par :

$$\delta_y(x) = \begin{cases} 1 & \text{si } x = y, \\ 0 & \text{sinon.} \end{cases}$$

L'équation 2 représente l'espérance, évaluée sur l'ensemble d'apprentissage, des densités régionales de luminosité, estimées empiriquement. Pour notre étude, nous disposons d'images injectées avec un produit de contraste, notées (+IV), et d'images non injectées, notées (-IV). Ces deux types d'images ont des plages de densités différentes, par conséquent, il est nécessaire de construire un modèle *a priori* pour chaque type de données. La figure 2 présente les distributions estimées sur une base de 67 images. Notons premièrement l'augmentation de la moyenne des niveaux de gris des images (+IV) due à l'injection du produit de contraste. Remarquons également que l'étendue de la distribution des images (+IV) est plus importante que celle des images (-IV). Ceci est dû essentiellement à l'hétérogénéité des niveaux de gris des régions des reins

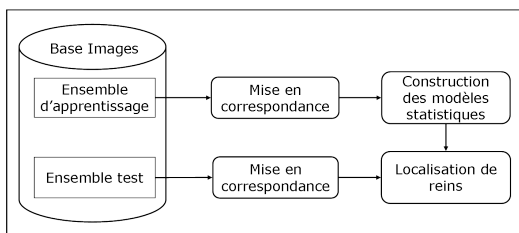


FIG. 1 – Organisation de la méthode de localisation des reins.

sur les images (+IV). En effet, nous disposons d'images acquises à différentes phases de la propagation du produit de contraste dans les tissus du rein, ce qui a pour effet d'augmenter l'hétérogénéité des tissus rénaux, et par conséquent la variance de la distribution des intensités observées.

3.2 Prétraitement des images

La construction d'un modèle spatial *a priori* des reins requiert un ensemble d'apprentissage constitué d'images bien recalées afin d'éviter au maximum les variations de position des différentes régions. Cette problématique relève du recalage d'images [18, 26, 29]. Notons qu'ici seules les régions des reins sont importantes pour la résolution du problème.

Le recalage d'image est un problème très difficile à résoudre et il l'est d'autant plus lorsque les données sont de nature médicale [18]. De nos jours, les performances des méthodes de recalage sont acceptables lorsque la transformation recherchée est globale (rigide, affine) ou quand le champ de déformation est très local (petits déplacements). Leur application reste donc problématique sur des images de l'abdomen et inter-sujet. La grande variabilité de leur contenu (présence de pathologies, morphologie du patient, niveau de coupe, etc.) et la nature molle des tissus rend cette tâche très délicate. En effet, l'hypothèse d'une correspondance point à point n'est manifestement pas réaliste et la recherche d'un homéomorphisme est une tâche illusoire [26]. Nous pouvons néanmoins citer les travaux de Park et al. [30]. Les auteurs effectuent un recalage élastique sur des images abdominales afin de construire un atlas probabiliste. Cependant, ce recalage est effectué sur chaque organe abdominal séparément, qui est détecté au préalable manuellement.

La construction de notre *a priori* spatial doit être précédée d'un recalage élastique des données. Cette étape est très délicate à réaliser automatiquement et nécessite que les deux images à recalcr ne soient pas trop différentes. Notons qu'il est souvent précédé d'un recalage rigide. Pour cette raison, nous avons jugé utile de commencer par tes-

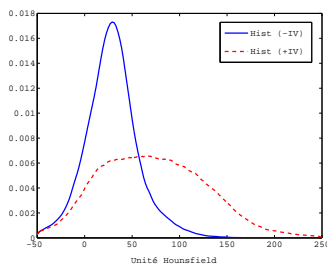


FIG. 2 – Modèle *a priori* des niveaux de gris pour les images sans produit de contraste (-IV) (courbe en trait continu) et pour les images avec produit (+IV) (courbe en trait discontinu).

ter les performances d'un recalage rigide. Notre choix s'est porté sur une méthode relativement récente basée sur la maximisation de l'information mutuelle par une approche multirésolution [37]. Nos tests ont confirmé nos craintes quant à la précision et la robustesse des résultats. En effet, les paramètres de la transformation sont souvent mal estimés (minimum local) en particulier dans les situations où le foie occupe un grand espace dans l'abdomen. L'utilisation d'un critère global est donc mal adaptée. Une des récentes directions de recherche pour pallier ce problème est l'utilisation des mesures de similarités locales [7, 26].

Comme alternative, nous avons développé un prétraitement complètement automatique afin de réduire la variabilité spatiale des régions des reins². Ce dernier s'est avéré relativement efficace et bien adapté à notre application. Pour la compréhension de la méthode proposée, quelques définitions s'imposent. Nous présentons à la figure 3 une coupe de l'abdomen, sur laquelle nous avons marqué les différentes parties de l'image. On appelle *champ de vue*, l'image du corps reconstruite à partir des projections des rayons X. Cette image est constituée d'une première couche fine de *peau* ensuite d'une couche de *graisse* et enfin de la *région abdominale*. On entend par *région abdominale* l'enveloppe englobant les différentes structures abdominales sans la peau et sans la couche de graisse. Le trait situé en bas de l'image sous le *champ de vue*, représente la table sur laquelle le patient est allongé. Notons que cette trace n'apparaît pas sur toutes les images.

L'observation des images de notre base de données (e.g. Fig. 5) nous a permis de constater que la variation de la position des reins dans l'abdomen est due essentiellement aux facteurs suivants :

- i) L'épaisseur de la couche de graisse se trouvant autour de la région abdominale, qui est spécifique à chaque patient.
- ii) La position du champ de vue dans l'image, qui dépend du processus d'acquisition.
- iii) La taille du champ de vue dans l'image, qui dépend de la physiologie du patient.
- iv) La variation de la résolution spatiale des différentes images de la base, qui s'étend de 0.5703 à 0.9746 mm par pixel.

Il apparaît évident que le prétraitement devra atténuer, voire éliminer, les effets de ces facteurs afin de réduire la variabilité de la position spatiale des reins. Pour ce faire, nous proposons une méthode qui commence par détecter la région abdominale dans chaque image. Puis, toutes les régions abdominales détectées sont rééchantillonnées pour qu'elles aient toutes la même dimension.

²Ce travail est développé indépendamment des travaux de Tsagan et al. [39]. Nous n'avons pris connaissance de ces travaux que depuis la publication de Lin et al. [25].

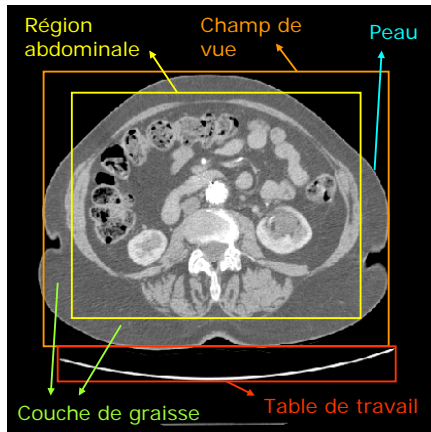


FIG. 3 – Un exemple d'image CT de l'abdomen avec une description des différentes parties la constituant.

3.2.1 Délimitation de la région abdominale

La première étape permet d'éliminer la couche de graisse ainsi que la fine couche de peau se trouvant autour. Les différentes phases de cette méthode sont détaillées ci-dessous :

- i) Un premier masque, appelé masque extérieur, est obtenu en appliquant un seuillage sur l'image. Le seuil est choisi égal à $-200 HU$, de telle sorte que toutes les structures abdominales et la table soient détectées. Une fois le masque créé, une opération d'ouverture, avec un disque de rayon 7 pixels, est appliquée afin d'éliminer la trace de la table. Ensuite, une opération d'érosion, avec un disque de rayon 10 pixels, est appliquée dans le but d'effacer, cette fois, la fine couche de peau. Le masque extérieur résultant est illustré par la figure 4(b).
- ii) La création d'un deuxième masque ne contenant pas les tissus graisseux, appelé masque intérieur, est obtenu avec une seconde opération de seuillage. Sur des images CT, la densité de la graisse varie entre -50 et $-100 HU$. Nous avons donc choisi un seuil égal à $-40 HU$. Les pixels isolés restant suite à cette opération de seuillage sont éliminés en utilisant une opération d'ouverture avec un élément structurant carré de taille 3×3 pixels. Le masque créé est présenté à la figure 4(c).
- iii) Finalement, les deux masques binaires résultant de ces deux opérations sont combinés avec un ET logique pour détecter la région abdominale, délimitant les différentes structures abdominales en éliminant la couche de graisse et la trace de la table. L'image résultante est présentée à la figure 4(d).

3.2.2 Redimensionnement

Les images dont nous disposons ont des résolutions spatiales différentes. Dans un premier temps, les régions abdominales détectées sont donc redimensionnées chacune

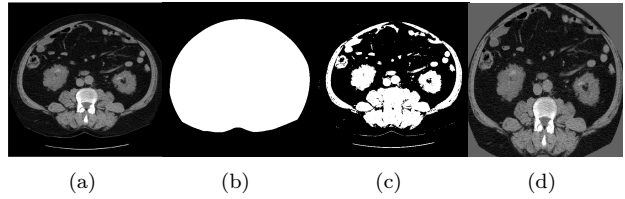


FIG. 4 – Détection de la région abdominale. a) Exemple d'une image CT ; b) Le masque extérieur créé pour éliminer la peau et la table de travail ; c) Le masque intérieur créé pour éliminer la couche de graisse sous la peau ; d) La région abdominale détectée.

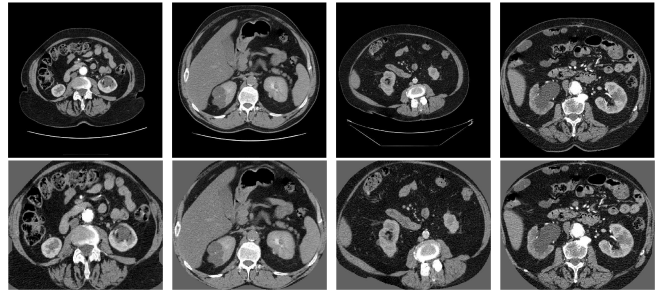


FIG. 5 – Exemples d'images CT et leurs régions abdominales détectées en utilisant la méthode de prétraitement proposée. 1^{ère} ligne) Images CT originales ; 2^{ème} ligne) Régions abdominales détectées.

à (dx^i, dy^i) , i étant le numéro de l'image, pour avoir la même résolution spatiale ($1mm^2$). Les régions ainsi obtenues sont ensuite redimensionnées pour avoir toutes les mêmes dimensions (\hat{dx}, \hat{dy}) que nous avons choisies d'approcher par la moyenne des (dx^i, dy^i) ; i.e.

$$(\hat{dx}, \hat{dy}) = (\mathbb{E}[dx^i], \mathbb{E}[dy^i])$$

où \mathbb{E} est l'espérance mathématique. Cela bien évidemment nécessite un rééchantillonnage des données. Nous avons jugé qu'une interpolation bilinéaire est suffisante pour notre application.

Quelques exemples d'images prétraitées sont montrés figure 5. Ce prétraitement est utilisé pour la construction des modèles spatiaux *a priori* pour les reins gauche et droit. Ces modèles sont présentés à la figure 6(b). La différence visuelle entre ces modèles et ceux construits en utilisant les images originales sans prétraitement (Fig. 6(a)) semble légère, mais quand ils sont utilisés pour la localisation des reins, cela aboutit à des taux de réussite très différents. Tous ces résultats sont présentés dans la section 5.

3.3 Algorithme de localisation des reins

Nous avons formulé le problème de la localisation des reins comme un problème d'optimisation d'un critère qui permet de positionner deux fenêtres, notées f_g et f_d , res-

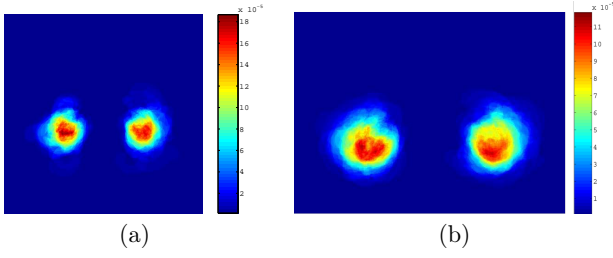


FIG. 6 – Modèles *a priori* spatiaux construits pour les reins gauche et droit en utilisant : a) les images sans prétraitement, b) les images prétraitées avec la méthode proposée.

pectivement sur les reins gauche et droit. L'appartenance d'un pixel s à une fenêtre f de support spatial $\Omega_f \subset \Omega$ est définie au sens statistique³, i.e. par une variable aléatoire $Z = \{z_s, s \in \Omega, z_s \in \mathbb{R}^2\}$, dont la densité de probabilité $p_Z(z_s)$, est caractérisée par :

$$\sum_{s \in \Omega} p_Z(z_s) = \sum_{s \in \Omega_f} p_Z(z_s) = 1$$

Le critère à optimiser prend en compte les modèles *a priori* de la position spatiale et des niveaux de gris d'une part et la propriété de symétrie des reins d'autre part. En effet, les reins se trouvent de part et d'autre d'une vertèbre et sont quasi symétriques. Cette symétrie sera caractérisée par une mesure de corrélation au sens statistique entre les niveaux de gris des deux fenêtres. Cette mesure est plus forte quand les deux fenêtres détectées contiennent des tissus de même nature. Nous proposons comme critère la fonction de coût suivante :

$$J(f_g, f_d, \Theta) = \frac{1}{2} \sum_{i \in \{g, d\}} \{E_{pdf}(f_i) + \lambda_1 E_{spat}(f_i)\} + \lambda_2 E_{sym}(f_g, f_d), \quad (3)$$

où $\lambda_1 \in \mathbb{R}^+$, $\lambda_2 \in \mathbb{R}^+$ sont des hyperparamètres, permettant de pondérer la contribution du terme spatial et du terme de symétrie dans la fonction de coût. Ce critère est minimisé par rapport à la variable $\Theta = \{s_g, s_d\}$ représentant respectivement les centres des fenêtres gauche f_g et droite f_d . Les différents termes qui le composent sont décrits ci-dessous.

3.3.1 Attache aux données, $E_{pdf}(f)$

Ce terme mesure la distance entre la distribution des niveaux de gris du modèle *a priori*, $p_\omega(l)$ et la distribution des niveaux de gris de la fenêtre f , notée $p_{\omega_f}(l)$, $l \in \mathcal{L}$, qui est définie par :

$$p_{\omega_f}(l) = \sum_{s \in \Omega_f} \delta_l(y_s) p_Z(z_s), \quad \forall l \in \mathcal{L}. \quad (4)$$

³L'indice f relatif à la fenêtre sera occasionnellement remplacé par g ou d dans les expressions où la distinction sera nécessaire, et ce pour toutes les variables relatives à la fenêtre. Nous noterons par exemple par Ω_g le support spatial de la fenêtre gauche.

La théorie sur les tests d'hypothèses nous est particulièrement utile pour le choix de la mesure de distance entre les deux distributions [12, 28]. Plusieurs mesures statistiques existent, nous en avons choisi deux : la divergence de Kullback-Leiber et la mesure du test de Kuiper [32]. E_{pdf} sera donc égale à l'une des deux expressions suivantes :

$$D_{KL}(p_{\omega_f} \parallel p_\omega) = \sum_{l \in \mathcal{L}} p_{\omega_f}(l) \log \frac{p_{\omega_f}(l)}{p_\omega(l)}, \quad (5)$$

$$D_{KP}(p_{\omega_f}, p_\omega) = \max_{l \in \mathcal{L}} [F_{\omega_f}(l) - F_\omega(l)] \quad (6)$$

$$+ \max_{l \in \mathcal{L}} [F_\omega(l) - F_{\omega_f}(l)] \quad (7)$$

où F_ω et F_{ω_f} sont les fonctions de répartition respectives aux distributions p_ω et p_{ω_f} .

3.3.2 Contrainte spatiale $E_{spat}(f)$

Soit $S = \{s \mid s \in \Omega, p(s) > 0\}$, où $p(s)$ est la probabilité du pixel s dans le modèle spatial *a priori* défini par l'équation 1. La probabilité que la fenêtre f soit incluse dans le rein (i.e. la probabilité que $\Omega_f \subset S$) est donnée par :

$$p(\Omega_f) = \sum_{s \in \Omega} p(s) p_Z(z_s) = \sum_{s \in \Omega_f} p(s) p_Z(z_s). \quad (8)$$

Nous définissons le terme spatial, au signe près, comme étant le logarithme de cette probabilité :

$$E_{spat}(f) = -\ln p(\Omega_f). \quad (9)$$

Ainsi, si $p(\Omega_f)$ tend vers 1, alors $E_{spat}(f)$ tend vers 0 et inversement plus la probabilité est faible, plus le terme E_{spat} devient important.

3.3.3 Symétrie gauche-droite $E_{sym}(f_g, f_d)$:

Le troisième terme permet de mesurer la similarité statistique entre les fenêtres gauche f_g et droite f_d . Pour ce faire, nous avons choisi parmi les mesures existantes, celles qui sont largement utilisées par la communauté de traitement d'images, à savoir l'information mutuelle et le coefficient de corrélation.

– **L'information mutuelle** : Elle mesure la dépendance statistique entre les deux fenêtres f_g et f_d . Elle est donnée par :

$$MI(f_g, f_d) = \sum_{l_g \in \mathcal{L}} \sum_{l_d \in \mathcal{L}} p(l_g, l_d) \log \frac{p(l_g, l_d)}{p_g(l_g) p_d(l_d)},$$

dans laquelle,

$$p(l_g, l_d) = \sum_{s \in \Omega_g} \delta_{l_g}(y_s) \delta_{l_d}(y_{\tilde{s}}) p_Z(z_s)$$

est la loi jointe et $p_g(\cdot)$ et $p_d(\cdot)$ sont les marginales respectives de f_g et f_d , \tilde{s} étant le pixel de la fenêtre

droite correspondant au pixel s de celle de gauche. Comme le rein gauche est la version miroir du rein droit, un retournement gauche droite est effectué sur l'une des fenêtres avant le calcul du terme de symétrie.

- **Le coefficient de corrélation :** Il mesure la dépendance linéaire entre les deux fenêtres. Il est défini par :

$$\rho(f_g, f_d) = \frac{\sum_{l_g \in \mathcal{L}} \sum_{l_d \in \mathcal{L}} l_g l_d p(l_g, l_d) - \mu_g \mu_d}{\sigma_g \sigma_d} .$$

Ici, μ_i et σ_i sont respectivement la moyenne et l'écart type de $p_g(\cdot)$ et $p_d(\cdot)$.

Nous cherchons à maximiser ces mesures ou encore à minimiser l'énergie suivante :

$$\begin{aligned} E_{sym}(f_g, f_d) &= -MI(f_g, f_d), \\ &\text{ou} \\ E_{sym}(f_g, f_d) &= -\rho(f_g, f_d) . \end{aligned}$$

3.4 Mise en oeuvre de l'algorithme d'optimisation

Comme la fonction de coût telle qu'elle a été définie auparavant n'est pas dérivable, nous nous sommes orientés vers des méthodes d'optimisation qui utilisent seulement l'évaluation de la fonction (i.e. sans aucune information du gradient). Nous avons utilisé la méthode de Nelder-Mead [22]. Elle permet d'optimiser les fonctions non linéaires à n variables en utilisant seulement les valeurs de la fonction, sans aucune information du gradient (implicite ou explicite). Chaque itération commence avec un simplexe à $(n + 1)$ sommets et les valeurs de la fonction évaluées en chaque sommet. Un ou plusieurs points candidats sont calculés et l'itération se termine par un nouveau simplexe, différent, de telle sorte que les valeurs de la fonction en ces sommets satisfassent une condition de descente comparée au simplexe précédent. L'algorithme de Nelder-Mead est particulièrement parcimonieux - comparé à d'autres méthodes de recherche directe - puisque, en pratique, une ou deux évaluations sont effectuées pour construire le nouveau simplexe.

Le problème majeur de ce type de méthode est l'initialisation. Pour cela, nous avons choisi d'initialiser l'algorithme avec M points tirés uniformément parmi ceux ayant une probabilité non nulle dans le modèle spatial *a priori* du rein (i.e. $p(s) > 0$). L'algorithme de minimisation converge en M minima locaux. Ensuite, le minimum optimal est sélectionné comme étant le minimum des minima locaux.

3.5 Choix de la forme de la fenêtre

Pour notre problème d'optimisation, nous fixons la même taille pour les deux fenêtres, nous cherchons donc à optimiser la fonction de coût par rapport aux centres de gravité s_g et s_d respectifs de deux fenêtres f_g et f_d .

a) Cas d'une fenêtre carrée :

La fenêtre f définie auparavant peut avoir n'importe quelle forme. Le cas le plus simple est celui d'un carré de taille fixe $(2n+1) \times (2n+1)$ avec une loi uniforme pour la densité de probabilité de la variable aléatoire Z :

$$p_Z(z_s) = \begin{cases} \frac{1}{\text{card}(\Omega_f)} & \text{si } s \in \Omega_f, \\ 0 & \text{sinon.} \end{cases}$$

Tous les pixels de la fenêtre ayant le même poids, ils influent donc de manière identique sur la fonction de coût $J(f_g, f_d)$.

b) Cas d'une fenêtre gaussienne :

Nous avons considéré également un cas plus général où la variable aléatoire Z suit une loi gaussienne centrée et de variance σ^2 , fixée au préalable :

$$p_Z(z_s) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\psi(z_s - z_{s_i})^2}{2\sigma^2}\right) , \quad (10)$$

où la fonction ψ mesure la distance spatiale entre le pixel s et s_i , le centre de gravité de la fenêtre f_i , $i \in \{g, d\}$. Ainsi la notion de fenêtre locale est progressive et elle est gouvernée par la valeur de σ .

4 Phase de délimitation du rein

Nous nous intéressons, dans cette section, à la seconde phase de l'approche d'identification automatique des reins. Cette phase permet la délimitation automatique des reins en utilisant les résultats de la phase de localisation. Le principe de notre approche consiste à déterminer un ensemble de régions candidates à partir de la fenêtre localisée sur le rein. Ensuite, un ensemble de règles sont apprises en se basant sur les descripteurs de ces régions afin de décider lesquelles appartiennent au rein. Le schéma global de cette phase est illustré par la figure 7. Ses différentes étapes sont décrites en détail dans les paragraphes suivants.

4.1 Segmentation

La segmentation est effectuée, dans un cadre bayésien, en utilisant le critère du maximum a posteriori (MAP) où les observations sont modélisées en utilisant un modèle de mélange de gaussiennes avec un *a priori* modélisé via un champ de Markov caché. En particulier, nous avons mis en oeuvre l'algorithme HMRP-Expectation-maximization, proposé by Zhang et al. [44] où l'estimation des paramètres du mélange est obtenue avec l'algorithme EM [13] et la minimisation de la fonction d'énergie est faite à l'aide

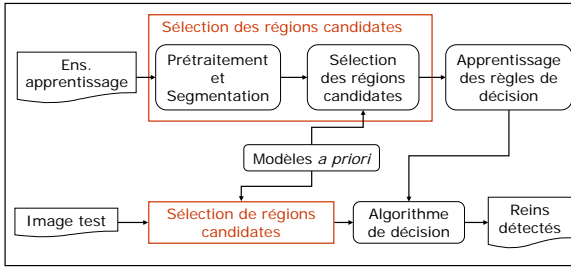


FIG. 7 – Architecture de la méthode de délimitation automatique des reins.

de l'algorithme déterministe ICM (Iterated Conditional Mode) [2]. Par ailleurs, la minimisation exploite une stratégie hiérarchique des données basée sur une combinaison d'un modèle multi-résolution (pour les observations) et un modèle multiéchelle (pour les étiquettes) [3].

La segmentation est complètement automatique. Toutes les images sont segmentées en un nombre fixe de classes (4 classes) et avec les mêmes paramètres de régularisation markovienne et de l'algorithme de minimisation hiérarchique. Une opération d'étiquetage est ensuite effectuée à l'aide de l'algorithme proposé par Suzuki et al. [36]. Cet algorithme, récent, a l'avantage d'être simple et rapide.

4.2 Sélection de régions candidates

L'idée ici, est de considérer les résultats de la première phase de localisation comme un point de départ pour déterminer l'ensemble de régions candidates à une étape de post-traitement. Nous faisons donc l'hypothèse que la fenêtre est bien localisée sur la région du rein. Par conséquent, il est naturel de considérer que toutes les régions ayant une intersection non nulle avec la fenêtre f appartiennent au rein. L'ensemble de ces régions est noté S_1 . Le rectangle englobant l'union des régions de l'ensemble S_1 , noté R_{box} , est ensuite déterminé. Un deuxième ensemble S_2 est défini, il est constitué de toutes les nouvelles régions ayant une intersection non nulle avec le rectangle R_{box} et une intersection nulle avec l'ensemble S_1 , i.e. toutes les régions de R_{box} n'appartenant pas à S_1 . Cet ensemble S_2 représente l'ensemble des régions *candidates* qui seront utilisées par la suite dans une étape de post-traitement. Celle-ci est une étape de classification supervisée permettant une identification automatique des régions appartenant au rein.

La figure 8 illustre l'algorithme de sélection des régions candidates. Pour une meilleure compréhension, nous présentons un exemple d'application de notre algorithme sur un prototype d'une partie de l'image de l'abdomen. Les résultats obtenus à chaque étape sont également affichés.

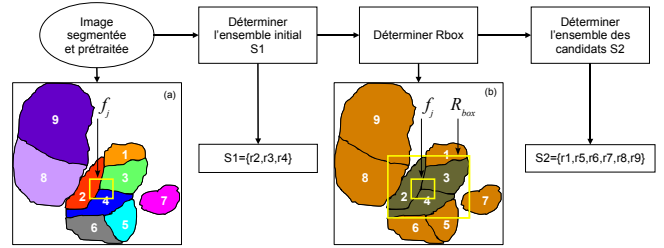


FIG. 8 – Algorithme de sélection des régions candidates. L'ensemble initial S_1 est déterminé en utilisant la fenêtre f_j localisée sur le rein de la figure (a). R_{box} est tracé sur la figure (b).

4.3 Calcul des attributs

Maintenant que l'ensemble des régions candidates est déterminé, nous procédons au calcul des attributs codant les caractéristiques de ces régions. Nous utilisons des attributs classiques de forme et de radiométrie tels que l'aire, le rectangle englobant ou la moyenne de la région considérée, et d'autres plus spécifiques à notre application. Ils sont plus pertinents car ils permettent de tirer profit des avantages apportés par l'introduction du modèle spatial *a priori* et du modèle *a priori* sur les niveaux de gris. En particulier, nous avons utilisé :

- Un attribut, noté $Spatial(r_i)$, qui représente l'énergie spatiale de la région r_i , calculée à l'aide de l'équation (9).
- Un attribut, noté $Pdf(r_i)$, qui représente l'énergie des niveaux de gris décrite à la section 3.3.1.
- Un attribut, noté $Vois(r_i)$, qui permet d'introduire une notion de voisinage entre les différentes régions candidates avec l'ensemble S_1 .

Ce dernier s'avère le plus discriminant. En effet, il est intuitif de penser que les régions les plus proches de S_1 peuvent appartenir au rein et que les régions les plus lointaines, voire non voisines, ont moins de chance d'appartenir au rein. Cette notion de voisinage est traduite par une mesure du degré de voisinage des régions candidates par rapport à S_1 . Nous avons défini le degré de voisinage d'une région r_i comme étant le plus court chemin⁴ entre cette région et l'ensemble S_1 .

Considérons d'abord, $S_2 = \{r_1, \dots, r_m\}$ l'ensemble des m régions candidates et considérons $R_{S_1} = \cup_{r_i \in S_1} r_i$, la région regroupant toutes les régions de l'ensemble initial S_1 . Le calcul de cet attribut utilise des outils de la théorie des graphes [17]. Après la construction de la matrice d'adjacence de l'ensemble des régions $A = \{R_{S_1}, r_1, \dots, r_m\}$, le plus court chemin est calculé avec, par exemple, l'algorithme de Dijkstra [14] entre toutes les régions de S_2 et R_{S_1} . Des résultats obtenus par application de cet algorithme sur l'exemple précédent sont fournis figure 9.

⁴Ici, la longueur du chemin entre deux régions r_1 et r_2 est définie comme étant le nombre d'arcs ayant un poids 1 qui existent entre la

4.4 Apprentissage de règles de décision

Nous cherchons à faire une classification binaire des régions candidates sélectionnées pour déterminer leur appartenance au rein. Vu la nature des données et des attributs des régions, il est intuitif d'essayer de classer de tels objets à travers une séquence de règles.

Nous avons développé une première méthode basée sur un ensemble de règles de décision *si - alors* dont les attributs, les seuils et l'ordre des tests sont déterminés d'une façon empirique. Les règles sont établies de manière à retranscrire la connaissance des experts. Les tests de cette méthode ont abouti à des résultats satisfaisants, suggérant que cette approche est adaptée pour résoudre notre problème de classification [4,38]. Cependant, son inconvénient majeur réside dans le choix empirique des attributs et des seuils, qui restent très dépendants des données. Cette méthode est donc peu robuste et très sensible au contenu de la base de données.

Un moyen efficace qui permet à la fois de renforcer la robustesse et de diminuer sa dépendance aux données, est d'utiliser des techniques d'apprentissage pour déterminer les règles et les seuils nécessaires. Parmi les techniques d'apprentissage existantes ayant le même mode opératoire que la méthode empirique proposée, les arbres de décision nous ont paru les mieux adaptés [5,33].

Les arbres de décision présentent quelques avantages par rapport à d'autres techniques de classification supervisée, en particulier dans le cas de données hétérogènes et de grande taille. Ils permettent une classification rapide en utilisant une séquence de règles simples et surtout la possibilité d'interpréter facilement la décision pour n'importe quel individu test, puisqu'elle est la conjonction de décisions prises tout au long du chemin parcouru, de la racine jusqu'à la feuille.

En particulier, nous nous sommes intéressés à la méthode d'arbre de décision binaire CART (Classification and Regression Tree) de Breiman et al. [5]. Cette approche est un partitionnement binaire et récursif des données ; elle favorise la construction d'un arbre de décision très grand

région r_1 et r_2 dans le graphe d'adjacence.

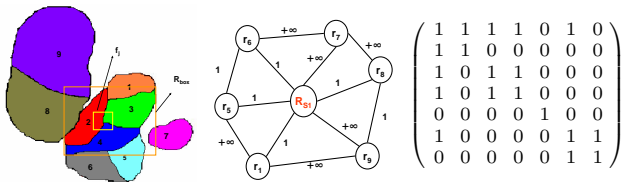


FIG. 9 – Exemple de calcul du degré de voisinage ; de gauche à droite : image prototype ; Graphe d'adjacence de l'ensemble $A = \{R_{S_1}, r_1, r_5, r_6, r_7, r_8, r_9\}$; Matrice d'adjacence et le degré d'adjacence de chaque région $r_i \in A$ par rapport à R_{S_1} est $Vois(A) = \{0, 1, 1, +\infty, 1, 2\}$. Ici $+\infty$ est utilisé quand les deux régions ne sont pas voisines.

TAB. 1 – Statistiques de la base d'images : âge des patients varie entre 19 et 88 ans, avec $\{55, 66, 74\}$ comme 1^{er}, 2^{ème} et 3^{ème} quartiles.

	Homme	Femme	+IV	-IV
# Patients	64	20	45	68
# Images	399	102	275	226

et procède ensuite à sa réduction d'une façon optimale.

L'algorithme CART procède en deux phases : une phase de construction ou d'expansion et une phase d'élagage. Pour la phase de construction, l'arbre est construit en divisant récursivement l'ensemble d'apprentissage jusqu'à ce qu'il devienne "pur", c'est à dire quand chaque partition contient des éléments de la même classe. L'idée fondamentale est de sélectionner chaque division d'un sous-ensemble de telle sorte que les sous-ensembles descendants sont plus "purs" que le sous-ensemble parent. La mesure de pureté s'effectue en utilisant une mesure de discrimination telle que l'entropie ou l'indice de Gini. Les parties les moins représentatives de l'arbre sont ensuite enlevées dans la phase d'élagage pour obtenir un arbre plus performant avec un plus fort pouvoir de généralisation.

5 Résultats

Dans le cadre de ce travail, nous avons collaboré avec le service radiologique du CHU d'Amiens pour l'étude des kystes des reins sur des images tomographiques. Elles sont acquises par différents radiologistes sur une période s'étalant sur une année. Dans certains cas, un produit de contraste est injecté et une à trois acquisitions sont effectuées en trois temps (à 30-40s, à 60-90s et à 90-120s). Par conséquent, nous trouvons dans la base de données considérée des images extraites de séquences avec produit de contraste et d'autres non injectées. Les images sont acquises avec des résolutions spatiales différentes (de 0.5703 to 0.9746 mm) et des pas intercoupe différents. Les statistiques calculées sur notre base de données constituée de plus de 500 images, représentatives de différents kystes, sont résumées dans le tableau 1.

Cette base d'images est divisée en un ensemble d'apprentissage (67 images), nécessaire pour la construction des modèles *a priori*, et un ensemble test (434 images), utilisé pour l'évaluation des méthodes de localisation et de délimitation des reins, et l'estimation des différents paramètres de celles-ci.

Le taux de succès est utilisé comme critère d'évaluation des performances des algorithmes proposés. Il est toujours calculé en fonction du nombre de reins et non en fonction du nombre d'images. Dans la phase de localisation, nous considérons comme un succès le cas où au moins les trois quarts de la surface de la fenêtre localisée appartiennent à la région du rein. Concernant la phase de délimitation, nous pensons qu'il est beaucoup plus important que la

région détectée contienne toutes les régions d'intérêts (rein + kystes) que l'inverse. Nous avons donc considéré comme succès le cas où au moins 85% du rein est détecté et au plus 150% de la surface du rectangle englobant le rein est détectée.

5.1 Approche de localisation des reins

La figure 10 présente quelques résultats obtenus en appliquant l'approche de localisation avec une fenêtre carrée sur les deux premières colonnes et avec une fenêtre gaussienne sur les troisième et quatrième colonnes. Apparaissent également sur cette figure : le cadre englobant la région abdominale, obtenue en appliquant la méthode de prétraitement proposée, et les petits carrés/cercles localisés automatiquement sur les reins.

5.1.1 Choix des paramètres de la fonction de coût

La méthode de localisation dépend des hyperparamètres λ_1, λ_2 , définissant les poids accordés respectivement au terme spatial et au terme de symétrie dans la fonction d'énergie, et de la taille de la fenêtre recherchée. Nous avons mené plusieurs expériences afin de déterminer la meilleure combinaison de paramètres. Ces tests sont effectués pour les quatre combinaisons possibles des mesures statistiques (Kuiper-CC⁵), (Kuiper-MI⁶), (Kullback-Leibler-CC) et (Kullback-Leibler-MI). La figure 11 nous montre quelques graphiques de la variation du taux de succès de localisation en fonction de la variation des hyperparamètres (λ_1, λ_2) et ce pour deux mesures du terme d'attache aux données : Kuiper et Kullback-Leibler. Les deux exemples montrés sont obtenus pour une fenêtre gaussienne isotrope avec un écart type spatial de 10 et l'information mutuelle pour l'estimation du terme de symétrie. Un comportement similaire est observé pour les autres

⁵CC désigne le coefficient de corrélation.

⁶MI désigne l'information mutuelle.

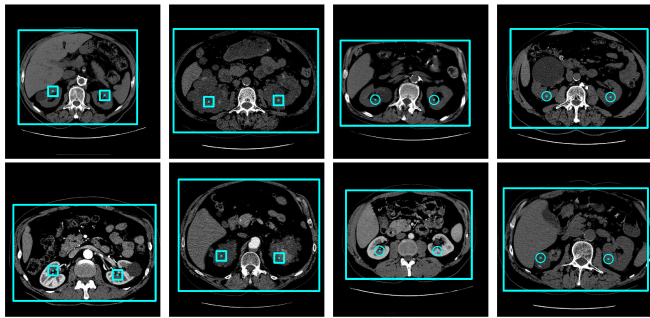


FIG. 10 – Exemples de localisation des reins dans le cas d'une fenêtre spatiale carrée avec $n = 16$ et dans le cas d'une fenêtre gaussienne avec $\sigma = 10$, sur des images (-IV) (1ère ligne) et des images (+IV) (2ème ligne). Toute les images possèdent un ou plusieurs kystes

combinaisons restantes et pour une fenêtre carrée.

L'observation de ces courbes nous permet de remarquer que : i) le taux de succès dépend étroitement du paramètre λ_1 , qui pondère la contribution de l'*a priori* spatial ; ii) le terme de symétrie influe faiblement sur les résultats de localisation ; iii) les résultats obtenus avec la mesure de Kuiper sont légèrement meilleurs que ceux obtenus en utilisant la divergence de Kullback-Leiber dans le calcul du terme d'attache aux données.

Nous avons également effectué des tests sur la variation du taux de succès pour les différentes combinaisons statistiques, en fonction de la taille de la fenêtre. Nous avons observé que les paramètres optimaux correspondent à une taille de (33×33) pour le cas carré (i.e $n = 16$) et un écart type $\sigma = 10$ pour le cas gaussien. Les meilleurs taux de succès pour les différentes combinaisons sont résumés au tableau 2. Nous remarquons que les performances de l'approche proposée sont légèrement meilleures sur les images sans produit de contraste. En effet, le produit de contraste accroît l'hétérogénéité de la région d'intérêt, ce qui se traduit par une augmentation de la difficulté du problème à résoudre. Par ailleurs, nous constatons une augmentation prévisible de l'importance de l'*a priori* spatial sur ce type d'image (λ_1 a toujours une valeur supérieure pour les images avec produit de contraste que sur celles sans produit).

5.1.2 Influence du nombre de points d'initialisation

Les résultats de l'algorithme d'optimisation de Nelder-Mead dépendent étroitement de l'initialisation. Pour cette raison, nous avons jugé nécessaire d'initialiser l'algorithme d'optimisation avec plusieurs points choisis aléatoirement de telle sorte que leurs probabilités sur le modèle spatial soient non nulles (i.e. $p(s) > 0$). Afin de mesurer la dépendance des résultats du choix de la procédure d'optimisation, une petite analyse statistique s'impose. Nous présentons en figure 12 l'évolution de la moyenne, du maximum et du minimum des taux de succès en fonction du nombre

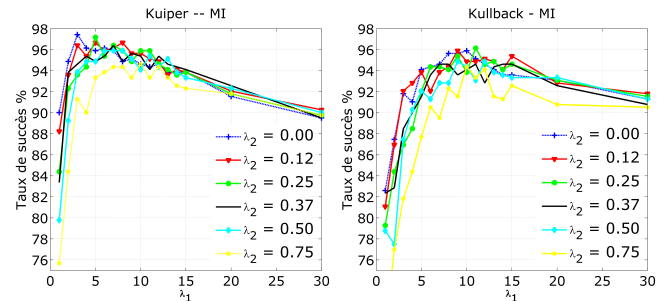


FIG. 11 – Évolution du taux de succès de la méthode de localisation du rein en fonction de l'hyperparamètre λ_1 pour différentes valeurs de λ_2 . Résultats obtenus pour une fenêtre spatiale gaussienne avec $\sigma = 10$.

TAB. 2 – Meilleurs taux de succès pour l'approche de localisation : comparaison des différentes combinaisons des mesures statistiques.

	# images	fenêtre	Test Kuiper	Test Kullback
(-IV)	196	carrée	97.44% MI, $n = 16$ $\lambda_1 = 6, \lambda_2 = 0.12$	94.88% MI, $n = 16$ $\lambda_1 = 10, \lambda_2 = 0.37$
		gauss	96.93% CC, $\sigma = 10$ $\lambda_1 = 4, \lambda_2 = 0.00$	95.40% MI, $\sigma = 10$ $\lambda_1 = 9, \lambda_2 = 0.12$
(IV)	238	carrée	95.13% MI, $n = 16$ $\lambda_1 = 10, \lambda_2 = 0.50$	91.74% MI, $n = 16$ $\lambda_1 = 12, \lambda_2 = 0.50$
		gauss	93.22% CC, $\sigma = 10$ $\lambda_1 = 12, \lambda_2 = 0.37$	92.37% MI, $\sigma = 10$ $\lambda_1 = 15, \lambda_2 = 0.25$

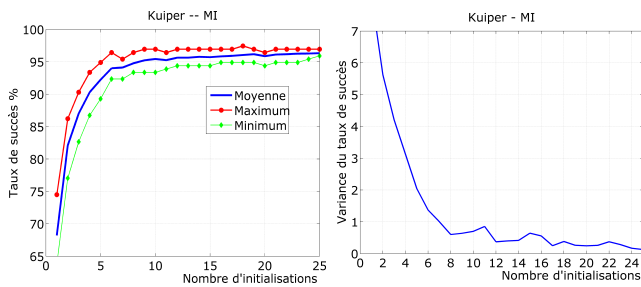


FIG. 12 – Influence du nombre de points d'initialisation de l'algorithme d'optimisation sur le taux de succès : variation du minimum, de la moyenne et du maximum (à gauche) et de la variance des taux de succès (à droite). Résultats estimés sur 20 expériences en utilisant une fenêtre gaussienne ($\sigma = 10$) et $(\lambda_1, \lambda_2) = (4, 0)$.

de points d'initialisation de l'algorithme. L'évolution de la variance du taux de succès est également montrée. Cette expérience suggère d'utiliser une dizaine de points pour obtenir une variance du taux de succès inférieure à 0.8. Ainsi, les résultats du tableau 2 doivent être arrondis à une décimale.

5.1.3 Importance du prétraitement

Un taux de succès de 82.6% seulement est obtenu si la méthode de localisation est testée sur les images originales ; i.e. sans l'étape de prétraitement. Cette statistique est obtenue sur les 196 images non injectées en utilisant la combinaison Kuiper-MI avec une fenêtre carrée. L'analyse du tableau 2 montre que le prétraitement proposé permet une amélioration de presque 18% des performances de la méthode de localisation. En effet, dans les mêmes conditions nous obtenons une performance de 97.4%. Ainsi, l'efficacité de la méthode de prétraitement proposée est bien évidente.

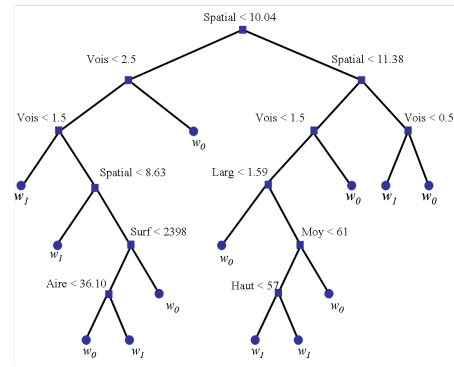


FIG. 13 – Arbre de décision optimal, construit avec la méthode CART en utilisant la validation croisée, pour la classification des régions du rein gauche sur des images sans produit de contraste. Ici, w_0 est la classe des régions n'appartenant pas au rein et w_1 est celle des régions appartenant au rein.

5.2 Approche de délimitation des reins

Comme décrit dans la section 4, nous avons utilisé la méthode CART pour construire un arbre de décision afin de classer les régions candidates et d'identifier celles appartenant à la région du rein. Par ailleurs, nous avons remarqué que les caractéristiques des régions se trouvant du côté gauche sont différentes de celles se trouvant du côté droit (e.g. les tailles des régions du côté droit sont plus importantes vu la présence du foie). Ainsi, il est préférable d'apprendre un arbre pour chaque côté. Outre le problème gauche/droit, nous sommes également confrontés au problème des images avec et sans produit de contraste. En effet ces deux types d'images n'ont pas les mêmes caractéristiques ; notamment en termes de statistiques des niveaux de gris. Par conséquent, nous avons jugé nécessaire d'apprendre quatre arbres de décision : deux pour les régions gauche et droite des images (-IV) et deux autres pour les images (+IV). La figure 13 présente un exemple d'arbre de décision, appris pour les images (-IV) et pour les reins gauches. Notons la simplicité de l'arbre, ce qui permet une classification assez rapide. Notons également le fort pouvoir de discrimination des attributs proposés, *Spatial* et *Vois*, spécifiques à notre application. Ces remarques sont également vraies pour le reste des arbres appris.

Nous présentons en figure 14 quelques exemples de détection des reins. Les fenêtres localisées en utilisant l'approche de localisation (petits carrés jaunes), la région abdominale détectée lors de l'étape de prétraitement (carrés rouges) ainsi que les régions des reins avec les kystes (courbes rouges) détectées automatiquement sont représentées.

Le tableau 3 présente les statistiques calculées en utilisant les arbres de décision sur les images (-IV) et (+IV). Nous illustrons les taux de succès de détection des reins

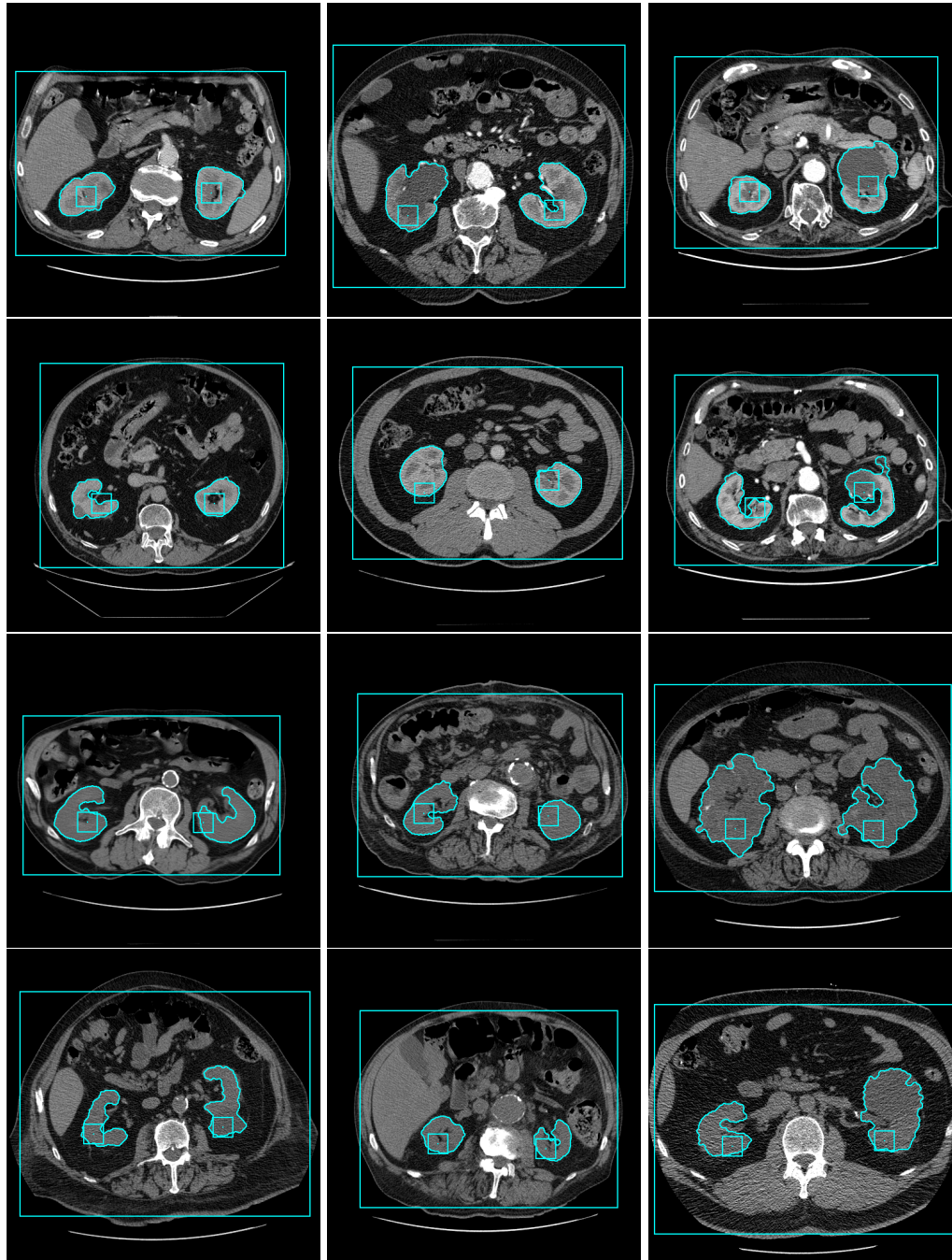


FIG. 14 – Exemples de résultats de délimitation des reins. Les deux premières lignes présentent des images avec produit de contraste et les deux dernières sans. Partant du haut vers le bas et ligne par ligne, les images 2, 3, 9 et 12 montrent des reins avec de gros kystes. L'image 4 montre l'exemple d'un petit kyste.

gauche et droit. Par ailleurs, les taux de Vrais Positifs (VP) et de Vrais Négatifs (VN), également présentés dans le même tableau, sont calculés afin d'apprécier la précision et la sensibilité relatives de cette méthode. Les taux de classification sont acceptables. Nous remarquons une performance légèrement supérieure de notre algorithme sur des images avec produit de contraste (+IV). Nous pensons que cette amélioration provient des meilleures performances de l'algorithme de segmentation sur ce type d'images. En effet, nous constatons que le taux le plus faible de l'algorithme de délimitation est obtenu pour des reins droit sur des image non injectées. Cela est dû essentiellement à des erreurs de segmentation du rein lorsque ce dernier est complètement collé au foie (e.g. fig. 15(d)). De telles erreurs de segmentation sont réduites voire éliminées par l'injection du produit de contraste; ce qui explique l'amélioration du taux de bonne délimitation.

5.2.1 Influence de la position initiale de la fenêtre localisée

L'approche proposée de délimitation des reins utilise les résultats de la première phase, la localisation. Selon la position de la fenêtre dans le rein, un ensemble de régions candidates est déterminé, il est ensuite classifié pour identifier les régions appartenant au rein. Ceci dit, on pourrait penser que les résultats de l'approche de délimitation des reins risquent d'être affectés par la position de cette fenêtre. Cependant, les expériences ont montré le contraire, heureusement d'ailleurs. En effet, afin de montrer statistiquement que les résultats de la méthode de délimitation des reins ne dépendent pas de la position de la fenêtre localisée sur le rein, nous avons testé notre approche indépendamment des résultats de la première phase. Ceci est possible en choisissant aléatoirement des fenêtres sur chaque rein (i.e. avoir des fenêtres de positions très variables) et de tester ensuite l'approche de délimitation des reins. Les résultats obtenus sont présentés au tableau 4. Nous remarquons que les taux moyens de bonne délimitation des reins restent dans le même ordre que ceux obtenus en utilisant les résultats de la phase de localisation (i.e. une variance très faible). Ceci nous permet de confirmer que l'approche proposée est très peu sensible à la position de la fenêtre sur le rein.

TAB. 3 – Taux de succès de l'algorithme de délimitation des reins en utilisant les résultats de l'étape de localisation. Ici les taux des vrais positifs et des vrais négatifs sont donnés pour le nombre de régions classifiées.

	-IV			+IV		
	VP%	VN%	Reins%	VP%	VN%	Reins%
Rein gauche	95.2	58.3	93.7	96.2	36.5	94.4
Rein droit	92.6	71.6	90.5	96.4	44.4	93.6

TAB. 4 – Taux de succès de l'algorithme de délimitation des reins en utilisant 10 initialisations aléatoires sur chaque rein (i.e. les résultats sont indépendants des ceux de la phase de localisation)

	-IV			+IV		
	VP%	VN%	Reins%	VP%	VN%	Reins%
Rein gauche	94.6	58.4	91.8	95.8	40.8	94.9
Rein droit	92.0	71.8	89.7	96.1	46.4	94.1

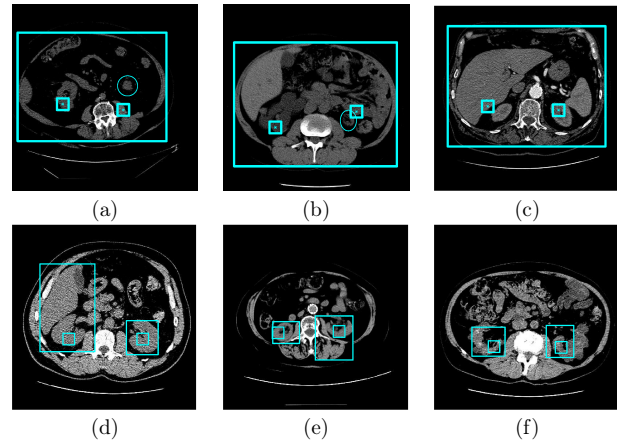


FIG. 15 – Quelques cas d'échec de l'approche d'identification des reins.

6 Discussion des cas d'échec

La figure 15 expose des exemples d'échec de notre approche d'identification. La première ligne montre des exemples de mauvaise localisation de la fenêtre sur le rein. Ceci est dû essentiellement à la taille minuscule des régions des reins (les vrais reins sont entourés par des cercles). La deuxième ligne montre des bonnes localisations des fenêtres sur les reins, cependant la phase de délimitation a échoué. Les deux premiers cas concernent la détection, en plus du rein, d'autres organes ou tissus avoisinant le rein. Ces échecs sont essentiellement dus à l'étape de segmentation. En effet, certaines images sont très bruitées et les organes adjacents ont des niveaux de gris et une texture identiques, ce qui ne permet pas une bonne séparation des tissus lors de la segmentation. Ceci engendre le problème de détection d'autres organes en plus du rein. Les seconds cas d'échec concernent la délimitation partielle des reins. Dans cette configuration, nous soulevons un problème qui porte sur la nature des kystes. Par exemple, la figure 15(f), présente un cas de polykystoses, une pathologie qui détériore tout le rein. Il est évident que dans ce cas, les régions segmentées seront fragmentées.

7 Conclusion

Une approche, complètement automatique, pour l'identification des reins avec des kystes sur des images scan-

ner 2D est proposée. Elle est constituée de deux phases : une phase de localisation qui utilise des modèles *a priori* spatial et des niveaux de gris, suivie d'une phase de délimitation de tout l'organe d'intérêt. La première phase prend la forme d'un problème d'optimisation d'un critère statistique et la seconde, d'un algorithme de classification supervisé de régions.

Nos expériences ont montré que les résultats obtenus sont satisfaisants d'autant plus que nous travaillons sur des images cliniquement acquises, donc particulièrement complexes. Nous sommes convaincus que la démarche proposée, en plus de son originalité, est adéquate au problème. Néanmoins, divers aspects de ce travail méritent sans doute d'être approfondis. Entre autre, la prise en compte d'information tridimensionnelle ne peut qu'améliorer le résultat. La résolution du problème en 2D nous a été imposé par l'équipe médicale du CHU d'Amiens, en vue de l'utilisation de cette approche dans un système de recherche images par le contenu destiné à la formation de médecins⁷. Les résultats ont été jugés concluant par l'équipe médicale. La poursuite des travaux dépend maintenant des priorités de financement établies par la région.

Remerciements

Nous remercions le professeur A. Remond et toute l'équipe médicale du service radiologique du CHU d'Amiens pour les données. Nous remercions le professeur Christophe Ambroise du Laboratoire 'Statistique et Génome' de l'université d'Évry pour les discussions sur les arbres de décision ainsi que les relecteurs de la revue *Traitement du Signal* pour leurs critiques constructives. Ce travail est financé dans le cadre d'un contrat de plan État-Région Picardie 2003-2009.

Références

- [1] N. Archip, P.J. Erard, M. Egmont-Petersen, J.M. Haefliger, and J.F. Germond. A knowledge-based approach to automatic detection of the spinal cord in CT images. *IEEE Trans. Med. Imag.*, 21(12) :1504–1516, December 2002.
- [2] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B*, 48(3) :259–302, 1986.
- [3] D. Boukerroui, A. Baskurt, J.A. Noble, and O. Basset. Segmentation of ultrasound images - multi-resolution 2D and 3D algorithm based on global and local statistics. *Pattern Recognition Letters*, 24(4-5) :779–790, February 2003.
- [4] Djamal Boukerroui, Wala Touhami, and Jean-Pierre Cocquerez. Automatic regions of interest identification and classification in CT images : Application to kidney cysts. In *First Workshops on Image Processing Theory, Tools and Applications*, pages 257–264, 2008.
- [5] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. *Classification And Regression Trees*. Belmont, CA, 1983.
- [6] O. Camara, O. Colliot, and I. Bloch. Computational modeling of thoracic and abdominal anatomy using spatial relationships for image segmentation. *Real-Time Imaging*, 10(4) :263 – 273, 2004.
- [7] Olivier Commowick, Vincent Arsigny, Aurélie Isambert, Jimena Costa, Frédéric Dhermain, François Bisdault, Pierre-Yves Bondiau, Nicholas Ayache, and Grégoire Malandain. An efficient locally affine framework for the smooth registration of anatomical structures. *Medical Image Analysis*, 12(4) :427–441, 2008.
- [8] T.T. Cootes, A. Hill, C.J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6) :355–366, July 1994.
- [9] T.T. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models – their training and application. *Computer Vision and Image Understanding*, 61(1) :38–59, January 1995.
- [10] D. Cremers, S.J. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69(3) :335–351, 2006.
- [11] Daniel Cremers, Mikael Rousson, and Rachid Deriche. A review of statistical approaches to level set segmentation : Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2) :195 – 215, April 2007.
- [12] R. D'Ágostino and M. Stephens. *Goodness-of-fit techniques*. Marcel Dekker, NY, USA, 1986.
- [13] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 34 :1–38, 1977.
- [14] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1 :269–271, 1959.
- [15] J. Duncan and N. Ayache. Medical image analysis : Progress over two decades and the challenges ahead. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(1) :85–106, 2000.
- [16] Igor Dydenko, Fadi Jamal, Olivier Bernard, Jan D'hooge, Isabelle E. Magnin, and Denis Friboulet. A level set framework with a shape and motion prior for

⁷Système de communication pour la formation en médecine basé sur l'indexation d'images'. Projet financé dans le cadre de l'article 29 du contrat de plan État-Région Picardie 2003-2009

- segmentation and region tracking in echocardiography. *Medical Image Analysis*, 10(2) :162–177, April 2006.
- [17] M. Gondran and M. Minoux. *Graphes et algorithmes*. Eyrolles, 1ère édition édition, 1995.
- [18] J. V. Hajnal, D.L.G. Hill, and D.J. Hawkes. *Medical image registration*. CRC press, June 2001.
- [19] T. Kaneko, Lixu Gu, and H. Fujimoto. Abdominal organ recognition using 3D mathematical morphology. In *Int. Conf. On Pattern Recognition*, pages 263–266, Barcelona, Spain, 2000.
- [20] L.J. Karssemeijer, N. van Erning and E.G. Eijkman. Recognition of organs in CT-image sequences : a model guided approach. *Comput Biomed Res.*, 21(5) :434–448, October 1988.
- [21] M. Kobashi and L.G. Shapiro. Knowledge-based organ identification from CT images. *Pattern Recognition*, 28(4) :475–491, 1995.
- [22] J.C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1) :112–147, 1998.
- [23] C.C. Lee, P.C. Chung, and H.M. Tsai. Identifying multiple abdominal organs from CT image series using a multimodule contextual neural network and spatial fuzzy rules. *IEEE Trans. Inf. Technol. Biomed.*, 7(3) :208–217, 2003.
- [24] M. Leventon, E. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proceedings of the IEEE Computer Vision on Pattern Recognition*, volume 1, pages 316–323, Hilton Head, SC, USA, Jun 2000. IEEE Computer Society.
- [25] D.-T. Lin, C.-C. Lei, and S.-W. Hung. Computer-aided kidney segmentation on abdominal CT images. *IEEE Trans. Inf. Technol. Biomed.*, 10(1) :59–65, 2006.
- [26] Grégoire Malandain. *Les mesures de similarité pour le recalage des images médicales*. Habilitation à diriger des recherches, Université Nice Sophia-Antipolis, March 2006.
- [27] A. Moreno, C.M. Takemura, O. Colliot, O. Camara, and I. Bloch. Using anatomical knowledge expressed as fuzzy constraints to segment the heart in CT images. *Pattern Recognition Letters*, 41(8) :2525–2540, 2008.
- [28] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 4 édition édition, 2001.
- [29] N. Paragios, Y. Chen, and O. Faugeras. *Handbook of Mathematical Models in Computer Vision*. Springer, (1st edition), October 31 2005.
- [30] H. Park, P.H. Bland, and C.R. Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans. Med. Imag.*, 22(4) :483–492, February 2003.
- [31] D.L. Pham, C. Xu, and J.L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2 :315–337, August 2000.
- [32] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C : The Art of Scientific Computing*. Cambridge University Press, NY, 1992.
- [33] R. Quinlan. *C4.5 : Programs for Machine Learning*. San Diego, 1993.
- [34] A. Shimizu, R. Ohno, T. Ikegami, H. Kobatake, S. Nawano, and D. Smutek. Segmentation of multiple organs in non-contrast 3D abdominal CT images. *Computer-Assisted Radiology and Surgery*, 2(3-4) :135–142, 2007.
- [35] Martin Spiegel, Dieter A. Hahn, Volker Daum, and Jakob. Segmentation of kidneys using a new active shape model generation technique based on non-rigid image registration. *Computerized Medical Imaging and Graphics*, 33(1) :29 – 39, 2009.
- [36] K. Suzuki, I. Horiba, and N. Sugie. Linear-time connected-component labeling based on sequential local operations. *Computer Vision and Image Understanding*, 89(1) :1–23, 2003.
- [37] P. Thévenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Trans. Image Process.*, 9(12) :2083–2099, December 2000.
- [38] Wala Touhami, Djamal Boukerroui, and Jean-Pierre Cocquerez. Fully automatic kidneys detection in 2D CT images : A statistical approach. In *Medical Image Computing and Computer-Assisted Intervention*, pages 262–269, 2005.
- [39] Baigalmaa Tsagaan, Akinobu Shimizu, Hidefumi Kobatake, and Kuniyoshi Miyakawa. An automated segmentation method of kidney using statistical information. In *Medical Image Computing and Computer-Assisted Intervention*, pages 556–563, 2002.
- [40] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky. Mutual information in coupled multi-shape model for medical image segmentation. *Medical Image Analysis*, 8(4) :429–445, December 2004.
- [41] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W.E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using levels sets. *IEEE Trans. Med. Imag.*, 22(2) :137–154, February 2003.
- [42] B. C. Vemuri, Y. Chen and J. Ye, and C. M. Leonard. Image registration via level-set motion : Applications

to atlas-based segmentation. *Medical Image Analysis*, 7(1) :1–20, March 2003.

- [43] J. Xie, Y. Jiang, and H.T. Tsui. Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Trans. Med. Imag.*, 24(1) :45–57, January 2005.
- [44] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.*, 20(1) :45–57, January 2001.



Djamal Boukerroui

Djamal Boukerroui est titulaire du diplôme d'ingénieur de l'École Nationale Polytechnique (ENP) d'Alger (1995), a obtenu le DEA SIP (Signal, Image, Parole) au CREATIS de l'INSA de Lyon (1996) et le grade de Docteur de l'INSA de Lyon en 2000. Après un post-doctorat de 30 mois au 'Medical Vision Laboratory' de l'Université d'Oxford, UK, il devient Maître de Conférences à l'Université de Technologie de Compiègne en septembre 2002 et intègre l'UMR 6599, Heudiasyc. Ses recherches portent sur la segmentation et la restauration d'images avec des applications en imagerie médicale et un intérêt particulier à l'imagerie ultrasonore.



Wala Touhami

Wala Touhami est née à Tunis (Tunisie) en 1977. Elle a obtenu le Dipl. Ingénieur en Génie Électrique à l'École Nationale d'ingénieurs de Tunis (ENIT) en 2001. Elle a eu son Doctorat en 2006 à l'Université de Technologie de Compiègne (France) avec une thèse sur l'identification et la classification des régions de reins kystiques sur des images tomographiques. Elle est Assistante à l'Institut Supérieur des Sciences Appliquées et de Technologie de Mateur (Tunisie). Elle est chercheur à l'ENIT, où elle travaille sur les problèmes de traitement d'images médicales, en particulier la segmentation et l'identification de région d'intérêt.



Jean Pierre Cocquerez

Jean Pierre Cocquerez est Professeur des Universités, il a reçu le grade de docteur d'état de l'université de Paris XI Orsay en 1981. De 1989 à 1999, il a été directeur de l'UMR 8051 CNRS ETIS de l'Université de Cergy Pontoise. Depuis 1999, il a rejoint l'Université de Technologie de Compiègne et il est membre du laboratoire HEUDIASYC UMR 6599. De 2001 à 2007, il a été directeur du département Génie Informatique de l'UTC. Depuis 2007, il est responsable du Master 'Sciences, Technologie, Santé'. Jean Pierre Cocquerez est aussi directeur du GDR 720 ISIS du CNRS depuis janvier 2006. Il est un relecteur régulier de IEEE ICIP et IEEE ICASSP. Ses domaines de recherche sont la segmentation d'images et la vision par ordinateur.

Phase based level set segmentation of ultrasound images

Phase-Based Level Set Segmentation of Ultrasound Images

Ahror Belaid, Djamal Boukerroui, Y. Maingourd, and Jean-Francois Lerallut, *Senior Member, IEEE*

Abstract—Ultrasonic image segmentation is a difficult problem due to speckle noise, low contrast, and local changes of intensity. Intensity-based methods do not perform particularly well on ultrasound images. However, it has been previously shown that these images respond well to local phase-based methods which are theoretically intensity invariant. Here, we use level set propagation to capture the left ventricle boundaries. The proposed approach uses a new speed term based on local phase and local orientation derived from the monogenic signal, which makes the algorithm robust to attenuation artifact. Furthermore, we use Cauchy kernels, as a better alternative to the commonly used log-Gabor, as pair of quadrature filters for the feature extraction. Results on synthetic and natural data show that the proposed method can robustly handle noise, and captures well the low contrast boundaries.

Index Terms—Echocardiography, level set, local phase, monogenic signal, segmentation.

I. INTRODUCTION

ULTRASOUND imaging is an exploration technique commonly used in many diagnostic and therapeutic applications. It has many advantages: it is noninvasive, provides images in real time, and requires lightweight material. However, ultrasound B-scan images are known to have low signal-to-noise ratio, low contrast, and high amounts of speckle [1]. This image texture, or speckle, is a correlated and multiplicative noise that inherently occurs in all types of coherent imaging systems. Hence, it makes modeling difficult as its statistics depend on the density and on the type of scatterers in the tissues [1]–[6]. All these characteristics make segmentation difficult, and therefore, complicate the diagnosis task.

A correct segmentation of structures is crucial in many medical applications. In clinical practice, the quantification of these structures is generally performed by manual tracing, which is a time consuming and by the application of geometrical assumptions that could introduce measurement errors in presence

of pathologies. Hence, reliable, rapid, accurate, and automatic or semiautomatic methods of structures extraction are required. Fully automatic segmentation of ultrasound images still remains a challenging topic [7].

Several approaches have been reported in the literature for automated or semiautomated border detection from ultrasound images. For instance, statistical models, arguing that these were more appropriate because of the significant noise and missing boundaries of ultrasound images, have been extensively used [7], [8]. For this reason, several probability density functions were used to model image gray levels statistics [9]–[18]. Probably, all exiting modeling paradigms and segmentation approaches have been tried on ultrasound data. A comprehensive recent survey is given in Noble and Boukerroui [7]. Recent works suggests that there is an increase of interest on statistical modeling of both uncompressed and compressed envelope of the backscattered signal, [8], and their use in level set segmentation algorithm [14], [16], [19]–[21]. Shape and time information are also of a great interest in a number of applications [16].

In this study, we refer to echocardiographic data. It is known that echocardiography has been one of the driving application areas of medical ultrasound and the literature on methods for automatically segmenting and tracking the left ventricle is extensive. As it has been pointed in [7], the most popular approach has been to treat echocardiographic endocardial segmentation as a contour finding approach. This is not straightforward as the contrast around the left ventricle chamber boundaries varies, depending on its relative orientation to the transducer direction, and to attenuation. Thus, conventional intensity gradient-based methods have had limited success on typical clinical images. To avoid this drawback, phase-based approach offers a good alternative, since it makes the approach robust to attenuation artifacts. It is within this framework that we propose an alternative in this paper.

A model of feature perception called the local energy model has been developed by Morrone *et al.* [22], [23]. This model postulates that features are perceived at points in an image, where the Fourier components are maximally in phase. A wide range of feature types give rise to points of high-phase congruency. These include step edges, line and roof edges, and Mach bands. It has been shown that this model successfully explains a number of psychophysical effects in human feature perception. Other important work on this model of feature perception can be found in [24]. Mulet-Parada and Noble [25], [26] were the first to successfully use the local phase information for boundary detection on echocardiographic images.

Some phase-based level set methods on segmentation [27]–[29], image enhancement [30], and registration [31], [32]

Manuscript received April 8, 2010; revised September 22, 2010; accepted October 25, 2010; Date of publication November 9, 2010; date of current version January 4, 2011. This work was supported by a grant of the Regional Council of Picardie and European Union/FEDER.

A. Belaid, D. Boukerroui and J.-F. Lerallut are with Heudiasyc UMR CNRS 6599, Université de Technologie de Compiègne, Centre de Recherche de Royallieu, 60205 Compiègne Cedex, France (e-mail: ahror.belaid@hds.utc.fr, djamal.boukerroui@hds.utc.fr, jean-francois.lerallut@hds.utc.fr).

Y. Maingourd is with the Pediatric Echocardiographic Center, Centre Hospitalier Universitaire Amiens, Amiens 80054, France (e-mail: maingourd.yves@chu-amiens.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITB.2010.2090889

on medical applications can be found in the literature. There is also a great increase of AM–FM methods and their use in medical applications [33], [34]. Also, phase information has been used in numerous applications, such as segmentation [35], wavelets [36] and AM–FM [35], [37], [38] image analysis, stereo matching [39], [40], optical flow [41], denoising [42], and corner and edge detection [24], [43]–[46]. Phase-based processing has attracted a lot of attention in image analysis, but probably not still enough in ultrasound image segmentation [7].

This paper concerns the development of a novel segmentation method of the left ventricle within the level set framework. This uses local phase information derived from the monogenic signal, which is a multidimensional extension of the analytic signal [44]–[46]. Our idea is to use a novel speed function, which combines the local phase and local orientation in order to detect boundaries in low contrast regions. A preliminary results of this work appeared in [47].

In the next section, we describe the extraction of local properties (phase, orientation, and amplitude) from 1-D and 2-D signals. The proposed segmentation method is presented in Section III. Section IV shows qualitative and quantitative experimental results on both synthetic and natural data. Section V provides a discussion followed by some concluding remarks in Section VI.

II. BACKGROUND

Openheimer and Lim [48] showed that when using two synthetic images, one of them containing only the phase information while setting its amplitude information to unity, and the other containing only the amplitude information while the phase is set to zero, only the image containing the phase is visible, although deteriorated. The image containing amplitude information is completely indiscernible. The information carried by the phase of a picture appears to be much more significant than the information carried by its amplitude. Indeed, the phase informs us about the location and orientation of image features, while the amplitude provides only information on their intensity. One of the popular methods to estimate local signal information is based on the analytic representation of the signal. Details for 1-D and n -D signals are given in the subsequent sections.

A. Monogenic Signal

To extract the local properties (amplitude and phase) of a 1-D signal $f(x)$, we need to represent it in its analytic form as presented in [49, Ch. 4]

$$f_A(x) = f(x) - i f_{\mathcal{H}}(x)$$

where $i = \sqrt{-1}$ and $f_{\mathcal{H}}(x)$ is the Hilbert transform of $f(x)$. The local amplitude (energy) and local phase of $f(x)$ are given by

$$A(x) = \|f_A(x)\| = \sqrt{f^2(x) + f_{\mathcal{H}}^2(x)},$$

$$\varphi(x) = \text{atan2}(f_{\mathcal{H}}(x), f(x)).$$

Recently, Felsberg and Sommer [45], [46], proposed a novel n -D generalization of the analytic signal based on the Riesz

transform, which is used instead of the Hilbert transform. Also, they proposed a 2-D isotropic analytic signal, called *monogenic signal*. This representation preserves the core properties of the 1-D analytic signal that decomposes a signal into information about structure (local phase) and energy (local amplitude). Felsberg and Sommer introduced the following filters in the frequency domain [44]:

$$H_1(u_1, u_2) = i \frac{u_1}{\sqrt{u_1^2 + u_2^2}}$$

$$H_2(u_1, u_2) = i \frac{u_2}{\sqrt{u_1^2 + u_2^2}}.$$

The spatial representation of the earlier filters is given by

$$h_1(x_1, x_2) = \frac{-x_1}{2\pi(x_1^2 + x_2^2)^{3/2}}$$

$$h_2(x_1, x_2) = \frac{-x_2}{2\pi(x_1^2 + x_2^2)^{3/2}}.$$

The monogenic signal \mathbf{f}_M is then defined as a 3-D vector formed by the signal $f(x_1, x_2)$ with its Riesz transform $\mathbf{f}_R = (\mathbf{h} * f)(x_1, x_2)$, with $\mathbf{h} = (h_1, h_2)$, by

$$\mathbf{f}_M(x_1, x_2) = (f, \mathbf{h} * f)(x_1, x_2) \quad (1)$$

$$= (f, \mathbf{f}_R)(x_1, x_2). \quad (2)$$

In the n -D case, the local phase is associated to a given local orientation due to the fact that structural information is related to a given orientation. The concept of *local phase vector* is proposed by Felsberg and Sommer [45], [46] and is defined by

$$\Phi = \varphi \cdot \mathbf{r} \quad (3)$$

$$\Phi = \text{atan}\left(\frac{\|\mathbf{f}_R\|}{f}\right) \frac{\mathbf{f}_R}{\|\mathbf{f}_R\|} \quad (4)$$

where $\text{atan}(\cdot) \in [0, \pi]$. The *monogenic phase* Φ (*local phase vector*), is similar to the product of 1-D local phase φ , multiplied by the orientation vector \mathbf{r} , if the underlying signal is *i* 1-D (intrinsically 1-D signals). The defined local phase vector can be interpreted as a rotation vector, which magnitude corresponds to the phase angle between the real signal and the monogenic signal, (see Fig. 1). In contrast to the 1-D case, the phase now includes additional geometric information. The monogenic phase completely characterizes the local gray-level transition of an image (i.e., local structure) as long as the image is locally *i* 1-D, given that the phase has been defined with respect to a given orientation. The phase vector orientation \mathbf{r} represents the local orientation of the image and is defined by the following relation [44]:

$$\theta = \text{atan}\left(\frac{h_2 * f}{h_1 * f}\right). \quad (5)$$

B. Quadrature Filters

In practical applications, the local properties are estimated using a pair of bandpass quadrature filters. Indeed, the detection of local properties by the monogenic signal assumes that the signal consists of few frequencies that is bandlimited. A real

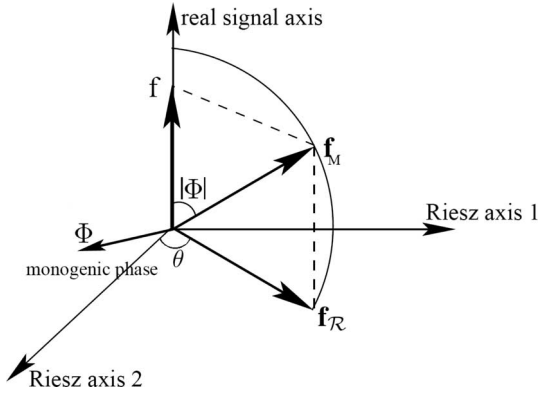


Fig. 1. Monogenic phase represented by means of a rotation vector. The amplitude of the rotation vector is the angle between the real and the monogenic signals. The rotation vector is orthogonal to the plane spanned by the real-signal axis and f_R .

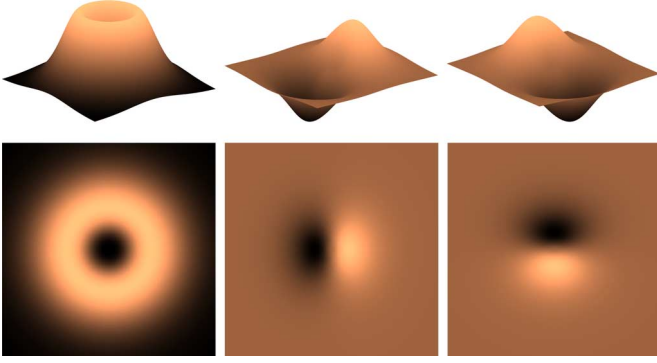


Fig. 2. 2-D Cauchy kernels in Fourier domain for a certain parameter $s > 0$. From left to right: isotropic even, and the pair of odd filters.

image consist of a wide range of frequencies, therefore a set of bandpass filters needs to be combined with the monogenic signal. Equations (1) and (2) become

$$f_M(x_1, x_2; s) = (c * f, c * h * f)(x_1, x_2) \quad (6)$$

$$= (c * f, c * f_R)(x_1, x_2) \quad (7)$$

where $c(x_1, x_2; s)$ is the spatial domain representation of an isotropic bandpass filter and $s > 0$ is a scaling parameter. Thus, the monogenic signal can be represented by a scalar-valued even and vector-valued odd filtered responses, with the following simple tick:

$$\text{even} = c * f$$

$$\text{odd} = (c * h_1 * f, c * h_2 * f).$$

Several families of quadrature pairs have been proposed and applied in the literature. Most authors have not provided a reasonable justification for the use of a particular family apart from simplicity of use or the satisfaction of the zero dc condition. In [50], [51], Boukerroui *et al.* compare several 1-D pairs of quadrature filters and concluded that log-Gabor kernels are probably not a very good choice in the case of feature detection. They showed that Cauchy family has better properties (see Fig. 2). In the frequency domain, a 2-D isotropic Cauchy kernel

is defined by

$$C(\mathbf{u}) = n_c |\mathbf{u}|^a \exp(-s|\mathbf{u}|), \quad a \geq 1 \quad (8)$$

where $\mathbf{u} = (u_1, u_2)$, s is a scaling parameter, and a/s is the peak tuning frequency of the filter. n_c is a normalization constant, see [50], [51] for more details.

In this paper, a Cauchy kernel is used as a bandpass filter. Interestingly, Cauchy filter could be seen as the a^{th} spatial derivative of the Poisson filter,¹ given in the frequency domain by

$$P(\mathbf{u}) = \exp(-2\pi|\mathbf{u}|s)$$

introduced in [46]. The authors have proved that the Poisson kernel establishes a linear and isotropic scale space: *the Poisson scale space*. The main property of this new scale space is its close relation to the monogenic signal. Hence, local phase and local amplitude, becomes inherent features of scale-space theory.

C. Edge Detection Measure

Step edge detection is performed using the *feature asymmetry* measure (FA) of Kovess [24] defined using the monogenic signal presented previously. To identify step edges essentially involves finding points, where the absolute value of the local phase is 0° at a positive edge and 180° at a negative edge. In other words, the difference between the odd and the even filter responses is large. Kovess suggested to use FA over a number of scales to detect step edge features. We define the multiple scales feature asymmetry

$$\text{FA} = \sum_s \frac{|\text{odd}_s| - |\text{even}_s| - T_s}{\sqrt{\text{even}_s^2 + \text{odd}_s^2} + \varepsilon} \quad (9)$$

where $|\cdot|$ denotes zeroing of negative values and T_s is the scale specific noise threshold [24]. The FA takes values in $[0, 1]$, close to zero in smooth regions and close to one near boundaries.

The application of this operator in [25] and [30] for ultrasound images has yielded good results. The authors used FA with steerable filter for boundary detection. However, it is expected to obtain better results using the monogenic signal, as it is the natural extension of the 1-D analytical signal.

III. DESCRIPTION OF THE MODEL

In this section, we present our proposed segmentation method. The idea is to use a novel-phase-based speed function within the level set framework.

Consider a gray level image as a function $I : \Omega \rightarrow \mathbb{R}^+$, where $\Omega \in \mathbb{R}^2$ is the image domain. The image gradient vector field is given by $\nabla I(x, y)$. Let us define the evolving contour $C : [0, L] \rightarrow \mathbb{R}^2$, given in a parametric form $C(p) = \{x(p), y(p)\}$, where p is an arclength parameter, and whose normal is defined by $\mathbf{n}(p) = \{-y_p(p), x_p(p)\}$.

The alignment term idea proposed by Kimmel *et al.* [52], is to search for a contour C that interacts with a given image, such

¹Here, the derivative is taken in the spatial domain in the radial direction, which preserves the isotropy property in higher dimension.

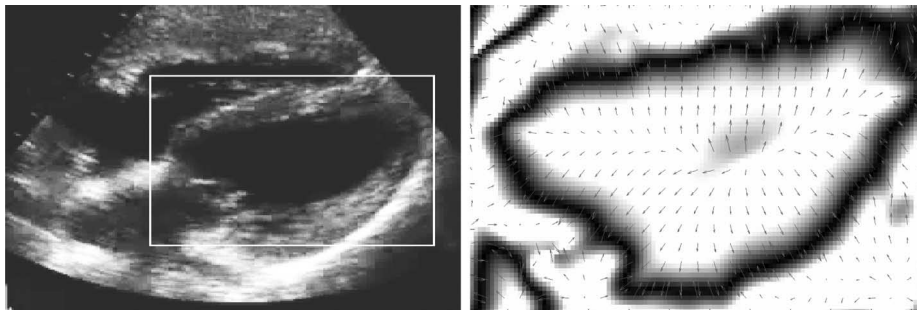


Fig. 3. (Left) Original echocardiographic image of a four chamber view with a defined region of interest (ROI). (Right) Underlying gray value image shows the g function given in (12) and the vector field show the local phase vector field, both calculated on the showed ROI.

that the curve's normal aligns with the gradient vector field. The alignment of the two vectors can be measured by their inner product that we denote by $\langle \mathbf{n}, \nabla I \rangle$. The geometric functional of the alignment measure in its robust form is given by [52], [53]

$$E_A(C) = - \int_0^L |\langle \mathbf{n}(p), \mathbf{V}(x(p), y(p)) \rangle| dp \quad (10)$$

where \mathbf{V} is a given vector field (e.g., image gradient). The functional $E_A(C)$ measures the alignment between the local image orientations and the curve's normals. In the proposed work, we use the local orientation computed by means of the monogenic phase Φ given by (4) instead of the classical gradient estimation. As shown in Fig. 3, the monogenic phase offers a good orientation estimation of the edge. The inner product $\langle \cdot, \cdot \rangle$ gets high values if the curve's normals align with the local orientations of the image. This is useful in low contrast boundaries.

In our active contour model, we will minimize the earlier alignment term and we will add some regularizing terms. It is known in literature that when we use several terms, geodesic active contour model (GAC) [54] serves as a good regularization for other dominant terms. Therefore, we introduce the following energy functional to be minimized:

$$E(C) = - \int_0^L |\langle \mathbf{n}(p), \mathbf{V}(x(p), y(p)) \rangle| dp + \lambda \int_0^L g(C(p)) dp + \nu \int \int_{\Omega_C} g(x, y) dx dy. \quad (11)$$

The second term is the GAC term, where g is an inverse edge indicator function, generally taken as $g(x, y) = 1/(1 + |\nabla G_\sigma * I|)$, and the integral is computed along the contour. Here, G_σ is the Gaussian kernel with standard deviation σ . The search in this case, would be for a curve along which the inverse edge indicator gets the smallest possible values. Thus, the values of g are close to one in smooth regions and close to zero near boundaries. The third term is a weighted region term as in [54]. It measures the area inside the curve C , i.e., the area of the region Ω_C .

In this paper, we define and use a phase-based edge indicator function instead of the classical inverse gradient based one, by the following new formulation:

$$g = 1 - (FA)^\alpha \quad (12)$$

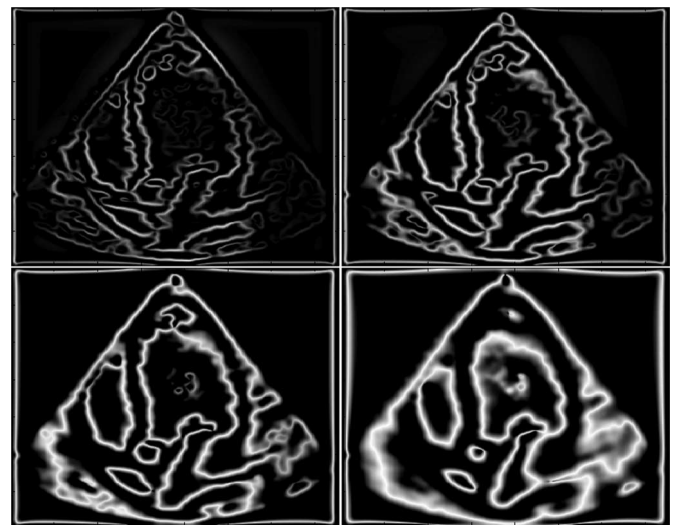


Fig. 4. Examples of feature asymmetry measure at different scales. From top-left to bottom-right: $s \in \{15, 20, 25, 30\}$, filters bandwidth = 2.5 octaves, i.e., $\alpha = 1.92$ (see [50]). The original image can be seen in Fig. 6 (left).

where $\alpha \in [0, 1]$ is a hyperparameter and $FA \in [0, 1]$ is the feature asymmetry measure, defined by (9). As it was mentioned in Section I, a recent works showed that ultrasound images respond well to phase-based edge detection. Moreover, a multiscales approach offers a better control on the edge detection quality. As shown in Fig. 4, by moving closer to finer scales, the FA measure recovers details and discontinuities, but loses regularity and continuity of the boundaries.

We embed a closed curve in a higher dimensional $\phi(x, y)$ function, which implicitly represents the curve C as a zero set, i.e., $C = \{(x, y) : \phi(x, y) = 0\}$. Following [53], the gradient descent flow minimizing (11), in the level set formulation, is given by

$$\frac{\partial \phi}{\partial t} = \left[-\mu \text{sign}(\langle \Phi, \nabla \phi \rangle) \text{div}(\Phi) + \lambda \text{div} \left(g(x, y) \frac{\nabla \phi}{|\nabla \phi|} \right) + \nu g(x, y) |\nabla \phi| \right] \quad (13)$$

where μ , λ , and ν are positives fixed parameters. The sign function in the first term can be simplified and written as $\text{sign}(\nabla_r \phi)$, where ∇_r is a gradient operator in the direction specified by \mathbf{r} , (see Appendix).

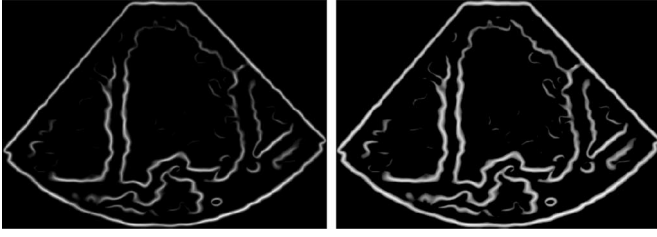


Fig. 5. Influence of the parameter α on the edge detection quality using the FA measure. (Left) $\alpha = 1$ and (right) $\alpha = 0.5$. The original image is given on Fig. 6 (right).

IV. EXPERIMENTAL RESULTS

MATLAB 7.6 (R2008a) was used for the implementation of the proposed method. Our program required approximately 15 s of CPU time per image (image size 256×256) on an IBM Intel Xeon single-CPU 3.4 GHz. In most of the experimental results shown in this section, the following parameters were fixed as such: bandwidth = 2.5 octaves as suggested in [50], wavelength = 20 pixels for natural data, and 10 pixels for synthetic data. These settings were roughly consistent with the ones used by Mulet-Parada and Noble [26]. We found by experiment that large α value would cause leakage of the endocardial contour. We therefore chose the largest α that did not have a leakage problem (see Fig. 5). In practice, we set $\alpha = 0.5$.

The GAC parameter λ is not set to the same value in all experiments. If we have to detect many objects of different sizes, then λ should be small. Conversely, if we have to detect only large objects, and to not detect smaller objects, then λ has to be larger. We also found by experiment that $\nu = 0.1$ was appropriate for most datasets.

The proposed variational level set method has been applied to a variety of natural and synthetic images. We compared the results of our approach with those of two other closely related algorithms. The first is the gradient-based version of our approach, which is the GAC presented in [54]. This algorithm has the same data adherence terms as our approach, but it lacks the alignment term. As suggested by Kimmel [55] for noisy images, the alignment term is turned off. The second algorithm (GAC + ML) is an alternative of the first one in which we strengthen it by a region-based term as presented in [14]. This model evolves according to image information using image gradient and an *a priori* knowledge about the statistical distribution of image gray levels. Specifically, the observed image gray levels are modeled by a Rayleigh distribution. The region-based term drives the curve evolution to achieve a *maximum likelihood* segmentation of the target, with respect to the statistical distribution law of image pixels.

Fig. 4 shows the edge detection function applied on real ultrasound image. It is computed by the FA measure, using the monogenic signal with the Cauchy filters. By moving closer to the coarse scales, edge detection looses details but recovers regularity of the boundaries. As shown in the top left of Fig. 4, finer scales do not detect weak boundaries, and may lead to segmentation leakages. Notice that the edges regularity can be controlled by the filter scaling parameter s and the GAC param-

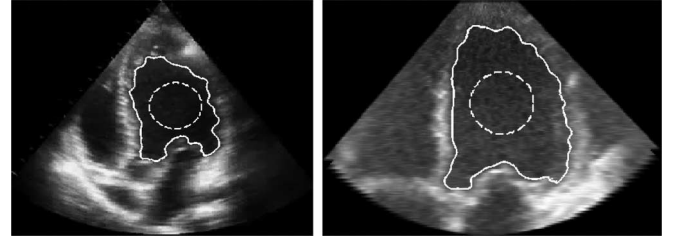


Fig. 6. Example of phase-based level set segmentation results of the left ventricle. The inner contours are the respective initializations.

eter λ . This is the reason why we have fixed the value of the scaling parameter and preferred adjusting the value of the GAC parameter in order to compare with similar approaches.

Fig. 6 shows illustrative results of our method on two typical ultrasound images (left ventricle). Fig. 7 shows illustrative comparison results of the proposed phase-based level set (PBLs) segmentation algorithm with the results of the GAC and the GAC + ML algorithms. These results give the reader some insight regarding the robustness to speckle noise and to attenuation.

In order to evaluate the proposed method and quantify its accuracy, we have used a set of manually delineated contours. We have collected a set of 20 bidimensional cardiac ultrasound images, obtained from a Philips IE33 echocardiographic imaging system. The dataset was segmented by two specialists in an independent way, i.e., in different days, at the Pediatric Echocardiographic Center, CHU Amiens. Each specialist segments each image five times, so that ten manual segmentations are available for each image. Thus, in all, we have 200 manual segmentations. This allows measuring the inter- and intraobserver variabilities, which are the differences performed by the same specialist as well as the differences between segmentations performed by different specialists, respectively.

Two distances have been used to compute the comparison between two contours. The first distance, referred here to as *dice similarity coefficient* (DSC), is given by [56]

$$\text{DSC}(S_1, S_2) = 2 \frac{|S_1 \cap S_2|}{|S_1| + |S_2|} \quad (14)$$

where S_1 and S_2 represent the ground truth and the obtained segmentation, and $|\cdot|$ denotes the cardinal of a set. The closer the DSC values to 1, the better is the segmentation. The second measure is the distance between two given curves, C_1 and C_2 represented as sets of ordered pairs of the x and y coordinates of points $C_1 = \{a_1, a_2, \dots, a_n\}$ and $C_2 = \{b_1, b_2, \dots, b_m\}$. The distance between the point a_i and the closest point on the curve C_2 is computed by

$$d(a_i, C_2) = \min_{b_j \in C_2} \|b_j - a_i\|.$$

These distances are computed for all points of the two curves and are averaged to yield the *mean absolute distance* between

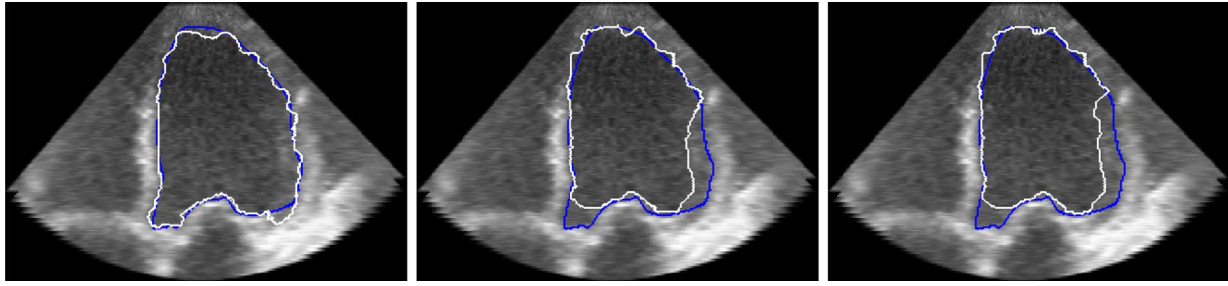


Fig. 7. Comparison of the PBLS (left), GAC (middle), and GAC+ML (right) results with a manual delineation. Blue line indicates manual delineation, and white line indicates semiautomatic segmentation.

TABLE I
PERFORMANCE INDICES MEASURES DSC (%) AND MAD (PIXELS) VERSUS
INTRA-OBSERVER, INTER-OBSERVER, AND COMPUTER-OBSERVER DISTANCES

Measures	DSC (%)			MAD		
	Mean	Median	SD	Mean	Median	SD
GAC	84.87	88.43	9.42	7.47	5.81	4.05
GAC+ML	86.78	89.08	8.85	6.61	5.19	4.32
PBLS	89.91	91.88	6.01	5.59	4.33	3.33
Intraobserver	96.15	96.51	0.92	2.39	2.20	0.65
Interobserver	93.66	93.82	3.67	3.80	3.06	2.48

Mean, median and standard deviation for DSC and MAD are shown.

the two curves $MAD(C_1, C_2)$ [57]. Hence,

$$MAD(C_1, C_2) = \frac{1}{2} \left[\frac{1}{n} \sum_{i=1}^n d(a_i, C_2) + \frac{1}{m} \sum_{i=1}^m d(b_i, C_1) \right]. \quad (15)$$

Table I shows a quantitative comparison between the three semiautomatic segmentation: our approach—that we denote in what follow PBLS for phase-based level set—, GAC, GAC + ML, and manual segmentation (intraobserver and interobserver) for the 20 echocardiographic images. The mean, median, and standard deviation of the 20 results are shown for both measures (see Fig. 8).

The first result is the segmentation performed by the classical-gradient-based GAC, the second one is the segmentation using GAC with the ML region term (see [14]) and the third one is the segmentation using our PBLS (11).

We have also used the simulation program Field II [58], [59], to synthesize phantom data with known ground truth. The phantom consists of 100 000 scatterers, and simulating 50 radio frequency lines. It consists of four columns with different contrasts, each one contains three circles with different scales (17, 20, and 23 pixels radius), see Fig. 10.

The segmentation results obtained by the GAC, GAC + ML model, and the proposed method are shown in Fig. 11. The results were obtained with $\lambda = 0.1$ for the GAC method and $\lambda = 0.6$ for PBLS and GAC + ML methods. Fig. 12 summarizes the quantitative evaluations of the three segmentation models on this dataset.

V. DISCUSSION

The quantitative evaluation on the natural data (see Tables I and II) show, as expected, that the use of GAC with an additional region term provides a significant improvement with respect to

the classical GAC. This is due to the fact that solely edge detection technique does not work for echocardiographic images with weak edges. The region term also improves the edge detection by avoiding local minima of the energy function, as a simple gradient descent is used. See, e.g., [19], [20] for an alternative minimization.

Now observe that the use of PBLS improves even more the already good results of the GAC + ML. Indeed, one of the major problems encountered in applying region-based method on ultrasound images, is the attenuation. In highly corrupted images with local intensity variations, the region term cannot segment the blood part of the left ventricle as a single region. Undoubtedly, the underlying assumption of a single tissue with Rayleigh statistics is not valid in such situations. The proposed phase-based term is more robust to attenuation artifact for being theoretically intensity invariant. It should also be remembered that in our approach, the phase-based GAC term is reinforced by the phase-based alignment term.

The intraobserver values are the mean of the values of DSC (%) and MAD after comparing all the manual segmentations of each specialist for each image. Interobserver values are the mean of the comparisons performed by different specialists. The interobserver results show that there is a significant difference between the segmentation of the two specialists (see details in Table II). As expected, the intraobserver differences are lower than the interobserver ones, and the later are less than all the semiautomatic ones (GAC, GAC + ML, and PBLS). The low variances of the intraobserver for both measures suggest a regularity of the segmentations of two specialists. This is confirmed by the details shown in Table II and the graphics shown in Fig. 8. Indeed, Fig. 8 shows separately the quality of segmentation associated to the specialists. In the same table and figure, we can detect images with low contrast and high speckle noise, which are difficult to segment by both physicians. For instance, image number 19 is difficult to segment, as it is suggested by the low mean and high standard deviation of the DSC measures for both physicians.

Table I suggests that the PBLS segmentations are not as good as the manual ones, but still very close to the ground truth. It can be observed, in Fig. 9, that the manual results are more regular, while the automatic results have more details. Note that our quantitative evaluation is similar to the one presented recently in [60]. The performances of our algorithm are of the same order as in [60].

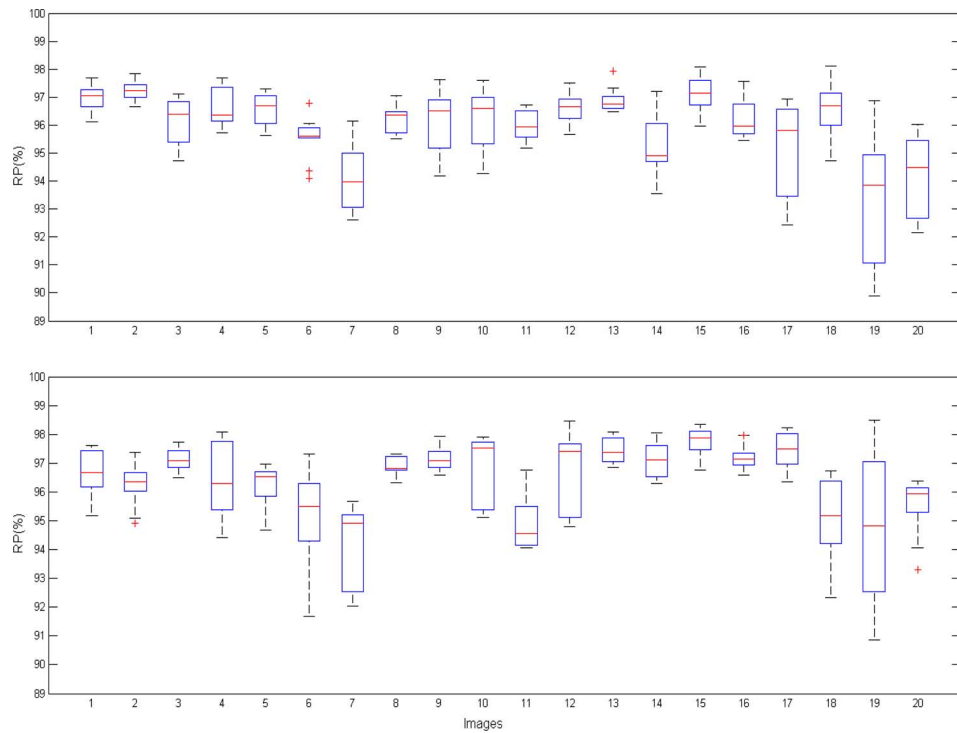


Fig. 8. Boxplots of the DSC (%) distance between all manual segmentations for the 20 images. (Top) Results of the first physician and (bottom) results of the second physician. The x -axis represents the image number.

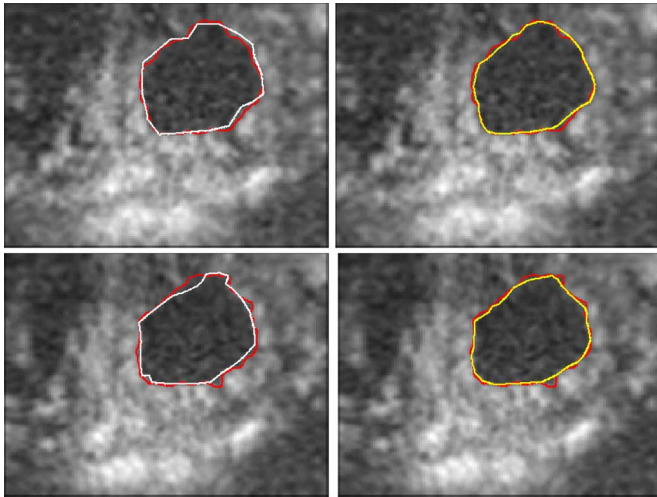


Fig. 9. Comparison of computer-generated segmentation (red) and the reference segmentation obtained from the manually delineated contours performed by the specialists. White in left column for the first specialist and yellow in right column for the second.

The experiments on the simulated data also confirm our observations regarding the relative performances of the three compared methods. The results of the proposed PBLs and GAC + ML methods are better than those of GAC. Although, the GAC provides also acceptable results as it can be observed in Fig. 11. The PBLs segmentation provides the best results in terms of av-

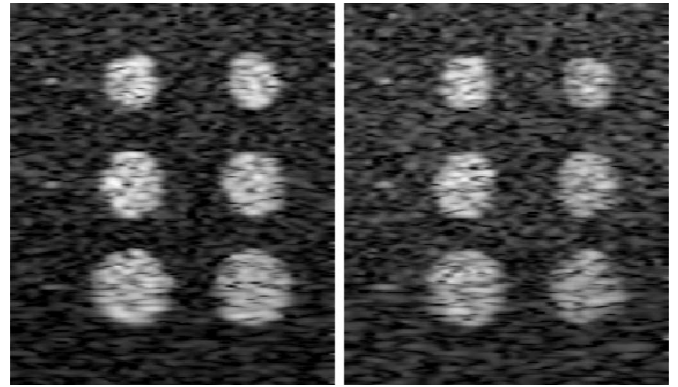


Fig. 10. Synthetic phantom. From left to right, high to low contrast circles. The three circles in each column, from top to bottom, have 17, 20, and 23 pixels radius, respectively.

erage performances and in terms of regularity as it has also the lower variances. An important observation is the large variance of the GAC + ML method. Our observation of its behavior suggests that it is mainly because of the simulated attenuation on some images that makes this variance increase. This reinforces our conclusion regarding the robustness of the proposed work relatively to other existing alternatives. Note, however, that recent works on region-based segmentation propose solutions for the segmentation of image with slow varying images statistics (see, e.g., [61] and [62]).

TABLE II
DSC (%) RESULTS OF THE TWO INTRA-OBSERVERS AND INTER-OBSERVER FOR TEN IMAGES

Images	1	2	3	4	5	6	7	8	9	10
Intraobserver 1	96.23	96.58	96.88	95.29	97.09	96.20	95.27	96.59	93.32	94.22
Intraobserver 2	96.74	96.71	97.43	97.07	97.78	97.22	97.41	95.07	94.74	95.48
Interobserver 1/2	87.61	89.27	97.17	96.96	97.13	95.96	87.61	87.24	93.30	92.82

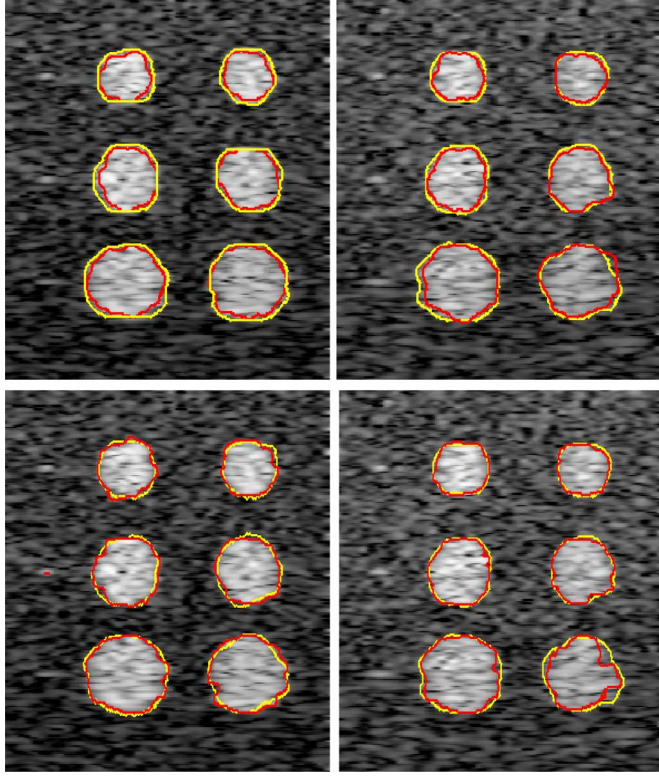


Fig. 11. Comparison of the PBLS segmentations (yellow) with the GAC segmentations (red, first line) and the GAC + ML segmentations (red, second line).

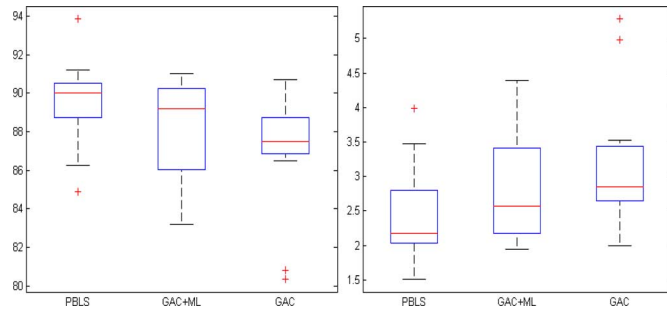


Fig. 12. Boxplot of the DSC (%) measure (left) and MAD distance (right) of the semiautomatic segmentation: PBLS, GAC + ML, and GAC.

VI. CONCLUSION

We have presented in this paper a new approach for the segmentation of the left ventricle in ultrasound images. In a level set framework, we integrate the use of a novel speed term based on local phase information and local orientation; both estimated using the monogenic signal. A key advantage of this approach is that it is more robust to intensity inhomogeneities.

The performance of the proposed PBLS segmentation is demonstrated on B-mode echocardiographic images and on re-

alistic synthesized images. Quantitative results are shown, when compared against hand-outlined boundaries, and are of the same order as in the state of art. Experiments on synthetic data show, as expected, that the PBLS approach outperform the classical-intensity-based GAC and GAC + ML models. Although the quantitative evaluation was performed on a limited dataset, our experiments show that the proposed terms can favorably replace classical ones of the same nature. It is interesting to investigate the addition of more recent region terms, as in [61] and [62], at least at the beginning of the segmentation process, in order to improve the capture range of the method.

APPENDIX

PHASE-BASED ALIGNMENT TERM

The Laplacian term presented in [52], in its implicit forme, is given by

$$\phi_t = -\text{sign}(\langle \nabla \phi, \mathbf{V} \rangle) \text{div}(\mathbf{V}) |\nabla \phi| \quad (16)$$

where ϕ is a level set function and \mathbf{V} is a vector field. The aim of our work is to build a Laplacian term based on a robust orientation estimation. Thus, we integrate the use of the phase vector Φ given in (4) as a vector field. The phase-based alignment term is then given by

$$\phi_t = -\text{sign}(\langle \nabla \phi, \Phi \rangle) \text{div}(\Phi) |\nabla \phi| \quad (17)$$

$$= -\text{sign}(\langle \nabla \phi, \varphi \cdot \mathbf{r} \rangle) \text{div}(\Phi) |\nabla \phi| \quad (18)$$

$$= -\text{sign}(\varphi \nabla \phi_r) \text{div}(\Phi) |\nabla \phi| \quad (19)$$

where ∇_r is the gradient operator in the direction specified by \mathbf{r} . Notice that, because $\varphi \in [0, \pi]$, the sign of the product $\varphi \nabla \phi_r$ depends only on the inner product between $\nabla \phi$ and the direction specified by \mathbf{r} . This leads to the following simplified formula:

$$\phi_t = -\text{sign}(\nabla \phi_r) \text{div}(\Phi) |\nabla \phi|. \quad (20)$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments, and Dr. Mathiron and Dr. Levy for their help in the clinical evaluation of the data.

REFERENCES

- [1] C. B. Burckhardt, "Speckle in ultrasound B-mode scans," *IEEE Trans. Sonics Ultrason.*, vol. SU-25, no. 1, pp. 1–6, Jan. 1978.
- [2] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Trans. Sonics Ultrason.*, vol. SU-30, no. 3, pp. 156–163, May 1983.
- [3] P. M. Shankar, J. M. Reid, H. Ortega, C. W. Piccoli, and B. B. Goldberg, "Use of non-Rayleigh statistics for the identification of tumors in ultrasonic B-scans of the breast," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 687–692, Dec. 1993.
- [4] V. Dutt and J. F. Greenleaf, "Ultrasound echo envelope analysis using a homodyned K-distribution signal model," *Ultrason. Imag.*, vol. 16, no. 4, pp. 265–287, Oct. 1994.
- [5] V. Dutt and J. F. Greenleaf, "Statistics of the log-compressed echo envelope," *J. Acous. Soc. Am.*, vol. 99, no. 6, pp. 3817–3825, Jun. 1996.
- [6] P. M. Shankar, "A general statistical model for ultrasonic backscattering from tissues," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 47, no. 3, pp. 727–737, May 2000.
- [7] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Jul. 2006.

- [8] S. Nadarajah, "Statistical distributions of potential interest in ultrasound speckle analysis," *Phys. Med. Biol.*, vol. 52, pp. 213–227, 2007.
- [9] D. A. N. Friedland, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Trans. Med. Imag.*, vol. 8, no. 4, pp. 344–355, Dec. 1989.
- [10] E. A. Ashton and K. J. Parker, "Multiple resolution bayesian segmentation of ultrasound images," *Ultrason. Imag.*, vol. 17, pp. 291–304, 1999.
- [11] M. Mignotte, J. Meunier, and J.-C. Tardif, "Endocardial boundary estimation and tracking in echocardiographic images using deformable template and markov random fields," *Pattern Anal. Appl.*, vol. 4, no. 4, pp. 256–271, 2001.
- [12] G. F. Xiao, M. Brady, J. A. Noble, and Y. Y. Zhang, "Segmentation of ultrasound B-mode images with intensity inhomogeneity correction," *IEEE Trans. Med. Imag.*, vol. 21, no. 1, pp. 48–57, Jan. 2002.
- [13] D. Boukerroui, A. Baskurt, J. A. Noble, and O. Basset, "Segmentation of ultrasound images: multiresolution 2D and 3D algorithm based on global and local statistics," *Pattern Recogn. Lett.*, vol. 24, no. 4–5, pp. 779–790, 2003.
- [14] A. Sarti, C. Corsi, E. Mazzini, and C. Lamberti, "Maximum likelihood segmentation of ultrasound images with rayleigh distribution," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 52, no. 6, pp. 947–960, Jun. 2005.
- [15] Z. Tao and H. Tagare, "Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images," *IEEE Trans. Med. Imag.*, vol. 25, no. 11, pp. 1483–1491, Nov. 2006.
- [16] Y. Zhu, X. Papademetris, A. J. Sinusas, and J. S. Duncan, "A coupled deformable model for tracking myocardial borders from real-time echocardiography using an incompressibility constraint," *Med. Image Anal.*, vol. 14, pp. 429–448, 2010.
- [17] T. Eltoft, "The rician inverse gaussian distribution: A new model for non-rayleigh signal amplitude statistics," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1722–1735, Nov. 2005.
- [18] T. Eltoft, "Modeling the amplitude statistics of ultrasonic images," *IEEE Trans. Med. Imag.*, vol. 25, no. 2, pp. 1722–1735, Feb. 2006.
- [19] Z. Tao and H. Tagare, "Tunneling descent level set segmentation of ultrasound images," in *Proc. Info. Med. Imag.*, 2005, pp. 750–761.
- [20] Z. Tao and H. D. Tagare, "Tunneling descent for m.a.p. active contours in ultrasound segmentation," *Med. Image Anal.*, vol. 11, no. 3, pp. 266–281, 2007.
- [21] M. Alemán-Flores, L. Álvarez, and V. Caselles, "Texture-oriented anisotropic filtering and geodesic active contours in breast tumor ultrasound segmentation," *J. Math. Imag. Vision*, vol. 28, no. 1, pp. 81–97, 2007.
- [22] M. C. Morrone, J. Ross, D. C. Burr, and R. Owens, "Mach bands are phase dependent," *Nature*, vol. 324, pp. 250–253, Nov. 1986.
- [23] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recogn. Lett.*, vol. 6, no. 5, pp. 303–313, 1987.
- [24] P. Kovsi, "Image features from phase congruency," *J. Comput. Vis. Res.*, vol. 1, no. 3, pp. 1–26, 1999.
- [25] M. Mulet-Parada and J. A. Noble, "2D+T acoustic boundary detection in echocardiography," in *MICCAI*. London, U.K.: Springer-Verlag, 1998, pp. 806–813.
- [26] M. Mulet-Parada and J. Noble, "2D+T acoustic boundary detection in echocardiography," *Med. Image Anal.*, vol. 4, no. 1, pp. 21–30, 2000.
- [27] I. Hacıhaliloglu, R. Abugharbieh, A. Hodgson, and R. Rohling, "Bone segmentation and fracture detection in ultrasound using 3d local phase features," in *MICCAI*, Berlin, Heidelberg: Springer-Verlag, 2008, pp. 287–295.
- [28] G. Låthén, J. Jonasson, and M. Borga, "Phase based level set segmentation of blood vessels," in *Proc. 19th ICPR*, Tampa, FL: IAPR, Dec. 2008, pp. 1–4.
- [29] R. Ali, M. J. Gooding, M. Christlieb, and J. M. Brady, "Phase-based segmentation of cells with brightfield microscopy," in *Proc. IEEE Symp. Biomed. Imag.*, 2007, pp. 57–60.
- [30] D. Boukerroui, J. A. Noble, M. C. Robini, and J. Brady, "Enhancement of contrast regions in sub-optimal ultrasound images with application to echocardiography," *Ultrason. Med. Biol.*, vol. 27, no. 12, pp. 1583–1594, 2001.
- [31] M. Mellor and M. Brady, "Phase mutual information as a similarity measure for registration," *Med. Image Anal.*, vol. 9, no. 4, pp. 330–343, 2005.
- [32] V. Grau, H. Becher, and J. A. Noble, "Registration of multiview real-time 3-D echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 26, no. 9, pp. 1154–1165, Sep. 2007.
- [33] V. Murray, S. E. Murillo, M. S. Pattichis, C. P. Loizou, C. S. Pattichis, E. Kyriacou, and A. Nicolaides, "An AM-FM model for motion estimation in atherosclerotic plaque videos," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2007, pp. 746–750.
- [34] V. Murray, E. S. Barriga, P. Soliz, and M. S. Pattichis, "Survey of AM-FM methods for applications on medical imaging," in *Proc. Ibero-American Conf. Trends Eng. Edu. Collab.*, Albuquerque, NM, Oct. 2009.
- [35] I. Kokkinos, G. Evangelopoulos, and P. Maragos, "Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 142–157, Jan. 2009.
- [36] M. Unser, D. Sage, and D. V. D. Ville, "Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2402–2418, Nov. 2009.
- [37] M. S. Pattichis and A. C. Bovik, "Analyzing image structure by multi-dimensional frequency modulation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 5, pp. 753–766, May 2007.
- [38] V. Murray, P. Rodriguez, and M. S. Pattichis, "Multiscale AM-FM demodulation and image reconstruction methods with improved accuracy," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1138–1152, May 2010.
- [39] D. J. Fleet, D. J. Fleet, A. D. Jepson, A. D. Jepson, M. R. M. Jenkin, and M. R. M. Jenkin, "Phase-based disparity measurement," *CVGIP: Image Understand.*, vol. 53, no. 2, pp. 198–210, 1991.
- [40] J. Zhou, Y. Xu, and X. Yang, "Quaternion wavelet phase based stereo matching for uncalibrated images," *Pattern Recogn. Lett.*, vol. 28, no. 12, pp. 1509–1522, 2007.
- [41] D. Zang, L. Wietzke, C. Schmaltz, and G. Sommer, "Dense optical flow estimation from the monogenic curvature tensor," in *SSVM*, 2007, pp. 239–250.
- [42] A. Wong, "Fast communication: Adaptive bilateral filtering of image signals using local phase characteristics," *Signal Process.*, vol. 88, no. 6, pp. 1615–1619, 2008.
- [43] P. Kovsi, "Phase congruency detects corners and edges," in *Proc. Aust. Pattern Recogn. Soc. Conf.: DICTA*, 2003, pp. 309–318.
- [44] M. Felsberg and G. Sommer, "A new extension of linear signal processing for estimating local properties and detecting features," in *Proc. 22nd DAGM Symp. Mustererkennung*, 2000, pp. 195–202.
- [45] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3136–3144, Dec. 2001.
- [46] M. Felsberg and G. Sommer, "The monogenic scale-space: A unifying approach to phase-based image processing in scale-space," *J. Math. Imag. Vis.*, vol. 21, no. 1, pp. 5–26, 2004.
- [47] A. Belaid, D. Boukerroui, Y. Maingourd, and J.-F. Lerallut, "Phase based level set segmentation of ultrasound images," in *Proc. 9th IEEE Int. Conf. Inf. Tech. App. Biomed.*, 2009.
- [48] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529–541, May 1981.
- [49] G. H. Granlund and H. Knutsson, *Signal Processing for Computer Vision*. Norwell, MA: Kluwer Publ., 1995.
- [50] D. Boukerroui, J. A. Noble, and M. Brady, "On the choice of band-pass quadrature filters," *J. Math. Imag. Vision*, vol. 21, no. 1, pp. 53–80, 2004.
- [51] D. Boukerroui, J. A. Noble, and M. Brady, *Frontiers in Robotics Research*, M. A. Denket Ed. New York: Nova Publ., 2006, ch. On the Selection of Band-Pass Quadrature Filters, pp. 67–111.
- [52] R. Kimmel and A. M. Bruckstein, "Regularized laplacian zero crossings as optimal edge integrators," *Int. J. Comp. Vision*, vol. 53, no. 3, pp. 225–243, 2003.
- [53] V. Caselles, R. Kimmel, and G. Sapiro, *Handbook of Image and Video Processing*, A. C. Bovik Ed., 2nd revised ed. New York: Academic, 2005, ch. Geometric active contours for image segmentation, pp. 613–629.
- [54] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comp. Vision*, vol. 22, no. 1, pp. 61–79, 1997.
- [55] R. Kimmel, "Fast edge integration," in *Geometric Level Set Methods in Imaging, Vision, and Graphics*. S. Osher and N. Paragios, Eds. New York: Springer, 2003, pp. 59–77.
- [56] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [57] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 15, no. 3, pp. 290–298, Jun. 1996.
- [58] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, vol. 39, no. 2, pp. 262–267, Mar. 1992.

- [59] J. A. Jensen, "Field: A program for simulating ultrasound systems," in *Proc. 10th Nordic-Baltic Conf. Biomed. Imag.*, 1996, vol. 34, pp. 351–353.
- [60] M. Alemán-Flores, L. Álvarez-León, and V. Caselles, "Texture-oriented anisotropic filtering and geodesic active contours in breast tumor ultrasound segmentation," *J. Math. Imag. Vision*, vol. 28, no. 1, pp. 81–97, 2007.
- [61] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2029–2039, Nov. 2008.
- [62] T. Brox and D. Cremers, "On local region models and a statistical interpretation of the piecewise smooth mumford-shah functional," *Int. J. Comp. Vision*, vol. 84, no. 2, pp. 184–193, 2009.

Ahrror Belaid received the M.S. degree in operation research from University of Abderrahmane Mira, Bejaia, Algeria. He is currently working toward the Ph.D. degree in the Department of Information Processing Engineering, Compiègne University of Technology, Compiègne, France.

His current research interests include the phase-based ultrasound images segmentation.



Djamal Boukerroui was born in Bejaïa, Algeria, in 1972. He received the B.S. degree in 1990 and the M.S. degree in electronics in July 1995 from Ecole Nationale Polytechnique of Algiers, Algeria, and the Ph.D degree in image processing from CREATIS Laboratory, INSA, Lyon, France, in 2000.

From March 2000 to August 2002, he was a Research Assistant at the Medical Vision Laboratory, Department of Engineering Science, University of Oxford, where he was involved on the analysis of echocardiographic image sequences. In September 2002, he joined the Department of Information Processing Engineering, Compiègne University of Technology, Compiègne, France as Maître de Conférences, where he is also a Member of HEUDIASYC [research unit associated with CNRS #6599]. His current research interests include low-level image processing, and specifically, its application in ultrasound image analysis.

Y. Maingourd's photograph and biography not available at the time of publication.



Jean-Francois Lerallut (SM'06) received the M.S. degree in electronics from Ecole Spéciale de Mécanique et d' Electricité, Paris, France, in 1976, and the Ph.D. degree from University of Technology of Compiègne (UTC), Compiègne, France, in 1979.

From 1979 to 2005, he was in the Department of Biomedical Engineering, UTC, where he is currently a Professor and also with the research unit HEUDIASYC associated with CNRS #6599.

Prof. Lerallut served in program committees and organizing committees for many international conferences. He is currently involved in projects on medical image processing.

Ultrasound image segmentation using local statistics with an adaptive scale selection

ULTRASOUND IMAGE SEGMENTATION USING LOCAL STATISTICS WITH AN ADAPTIVE SCALE SELECTION

Qing YANG and Djamel BOUKERROUI

Université de Technologie de Compiègne, Laboratoire Heudiasyc UMR-CNRS 7253
BP 20529 - 60205 Compiègne Cedex, France

ABSTRACT

We propose a region-based segmentation method based on local statistics. The adaptive spatial locality is defined using the Intersection of Confidence Intervals (ICI) approach. This pixel dependent local scale is estimated, conditionally on the current segmentation, in the sense of minimizing the mean-square error of a Local Polynomials Approximation (LPA). In other words, the scale is ‘optimal’ since it gives the best trade-off between the bias and the variance of the estimates. We provide a comparison with the single scale local region-based model. Results on simulated and real ultrasound images show that the proposed adaptive scale selection gives a robust solution to the attenuation problem.

Index Terms— local region statistics, Intersection of Confidence Intervals, ultrasound image segmentation.

1. INTRODUCTION

The use of statistical models in region-based image segmentation has a long tradition. Its introduction in active contour segmentation methods, mostly within the level set framework, has led to a considerable improvement in efficiency and robustness. For instance, the Chan-Vese (CV) model [1] and its variant [2] both consider image background and foreground as constant intensities represented by their mean values. The mean separation method [3] relies on the assumption that foreground and background should have maximally different intensities. These methods have many advantages and perform better than edge-based models in handling the noise and weak boundaries. However, they cannot deal with the intensity inhomogeneities, which is almost unavoidable.

Recently, some work has been carried out in utilizing local image statistics within the level set paradigm. For instance, a local binary fitting method has been proposed in [4]. A general framework for local region-based segmentation models has been presented in [5]. An interpretation of the piecewise smooth Mumford-Shah (MS) functional using local Gaussian models has been proposed in [6]. A local Gaussian distribution fitting energy is defined for the region-based active contour model in [7]. All of the above methods prove that the segmentation using local statistics has the ability to capture the boundaries of inhomogeneous objects.

Local region-based segmentation models, however, are found to be more sensitive to noise than global ones. Such models may also be more sensitive to initialization if the local scale is not appropriate. This brings out several problems that need to be addressed such as: Can global and local statistics be combined in one model? Is it possible to define a pixel dependent local scale from the estimation of image statistics? A first attempt within the level set framework has been proposed in [7]. Their approach is straightforward: it adds two energy functions of the same nature, where the model parameters are estimated globally in one and locally in the other. In fact, this is not the first time that global and local statistics are combined to solve a segmentation problem. To our knowledge, it is within the Bayesian framework that the first proposition has been introduced [8, 9].

This paper involves ideas described in [10] and applied to ultrasound (US) images. We propose a segmentation method based on local statistics with an adaptive local region, whose pixel dependent scale meets the counterbalance between the bias and the variance of estimations. Our method provides more desirable segmentation results compared to the single scale local region-based method [6].

2. LOCAL REGION-BASED SEGMENTATION

Recently, Brox and Cremers [6] have derived a statistical interpretation of the full MS functional [11] by relating it to recent works on local region statistics. They showed that the minimization of the piecewise smooth MS functional is equivalent to a first order approximation of a Bayesian a-posteriori maximization, basing on local region statistics. Thus this link is a consequence of the fact that local region models are not bounded by the strong assumption of identically distributed pixel intensities. Precisely, it is the approximation of the Bayesian setting with an additive noise of local Gaussian distribution, which can be expressed by the minimization of:

$$E(\phi) = \sum_r \int_{\Omega_r} \left[\frac{(I(\mathbf{x}) - \mu_r(\mathbf{x}))^2}{2\sigma_r^2(\mathbf{x})} + \frac{1}{2} \log(2\pi\sigma_r^2(\mathbf{x})) \right] d\mathbf{x} \quad (1)$$

$$+ \lambda \int_{\Omega} \delta(\phi(\mathbf{x})) |\nabla \phi(\mathbf{x})| d\mathbf{x} + \nu \int_{\Omega} H(\phi(\mathbf{x})) d\mathbf{x},$$

where the subscript $r = \{i, o\}$ represent inside and outside of the segmentation contour C , and \mathbf{x} is the spatial position in $\Omega \subset \mathbb{R}^2$. The image I is divided into foreground Ω_i and

background Ω_o by C . H is the Heaviside function of the level set function ϕ . The last two terms form the regularization term, whose contribution does not depend on image statistics. In Eq. (1), the local means and variances are functions of \mathbf{x} , and can be estimated using normalized convolutions:

$$\begin{aligned}\mu_r(\mathbf{x}) &= \frac{\int_{\Omega} K_h(\mathbf{x} - \zeta) H_r(\phi(\zeta)) I(\zeta) d\zeta}{\int_{\Omega} K_h(\mathbf{x} - \zeta) H_r(\phi(\zeta)) d\zeta}, \\ \sigma_r^2(\mathbf{x}) &= \frac{\int_{\Omega} K_h(\mathbf{x} - \zeta) H_r(\phi(\zeta)) (I(\zeta) - \mu_r(\mathbf{x}))^2 d\zeta}{\int_{\Omega} K_h(\mathbf{x} - \zeta) H_r(\phi(\zeta)) d\zeta},\end{aligned}\quad (2)$$

where $H_i = H(\phi(\zeta))$, $H_o = 1 - H(\phi(\zeta))$. $K_h(\cdot)$ can be any appropriate local kernel (often chosen to be the Gaussian kernel with a scaling parameter h). The minimization of Eq. (1) is obtained when each point on C has moved, such that the local interior and local exterior of each point along C are best approximated by local means $\mu_r(\mathbf{x})$. A fast implementation using recursive filtering is given in [6] and references therein.

3. ADAPTIVE WINDOW SIZE BASED ON LPA-ICI

Suppose f is a function of the spatial position \mathbf{x} . We wish to reconstruct f using its noisy observations $I(\mathbf{x}_s) = f(\mathbf{x}_s) + \varepsilon_s$, $s = 1, \dots, n$. The observations coordinates \mathbf{x}_s are known, and ε_s are random errors with variance σ^2 . Assume f can be approximated locally by members of a simple class of parametric functions. The LPA provides estimations in a point-wise manner, which finds the weighted least-square fitting in a sliding window. The estimate given by the LPA can be written as the kernel operator on the observations [12]:

$$\hat{f}_h(\mathbf{x}) = \sum_s g_h(\mathbf{x}, \mathbf{x}_s) I(\mathbf{x}_s), \quad (3)$$

where the kernel g_h is defined by the window w_h and the set of polynomials ψ_h . When the grid is assumed to be regular, the kernel $g_h(\mathbf{x}, \mathbf{x}_s)$ become shift-invariant on \mathbf{x} and the solution is given by a convolution operation.

The ICI rule aims at searching for the largest local window size (minimizing local variances) where the LPA fits well to the observations (minimizing local bias). Let \mathbf{h} be a set of the ordered scale values $\mathbf{h} = \{h_1 < h_2 < \dots < h_J\}$. The estimates $\hat{f}_h(\mathbf{x})$ are calculated for $h \in \mathbf{h}$ and compared. The ICI rule, which uses the estimates and their variances, identifies a scale closest to the ideal one. First a set of confidence intervals $Q(h)$ of the estimate is obtained [12]:

$$Q(h) = [\hat{f}_h(\mathbf{x}) - \Gamma \cdot \sigma_{\hat{f}_h}(\mathbf{x}, h), \hat{f}_h(\mathbf{x}, h) + \Gamma \cdot \sigma_{\hat{f}_h}(\mathbf{x}, h)]. \quad (4)$$

Here, $\sigma_{\hat{f}_h}^2(\mathbf{x}, h) = \sigma^2 \sum_s g_h^2(\mathbf{x} - \mathbf{x}_s)$ is the standard deviation of the estimates, which can be estimated given the noise model and its variance σ^2 . Eq. (4) is equivalent to: $\forall h_i \leq h^*(\mathbf{x}), f_{h_i}(\mathbf{x}) \in Q(h_i)$ holds with probability p , related to the threshold Γ . Thus, for all $h_i < h^*$, the $Q(h_i)$ have a point in common, namely $f(\mathbf{x})$. If the ICI is empty, it indicates $h_i > h^*$. In this way, the ICI rule can test the existence of this common point and to obtain the adaptive scale. The ICI algorithm, which are point-wise repeated for each \mathbf{x} , is defined by the following steps:

1. Define a sequence of confidence intervals $Q_i = Q(h_i)$ with their lower bounds L_i and upper bounds U_i :

$$L_i = \hat{f}_{h_i}(\mathbf{x}) - \Gamma \cdot \sigma_{\hat{f}_{h_i}}(\mathbf{x}, h_i),$$

$$U_i = \hat{f}_{h_i}(\mathbf{x}) + \Gamma \cdot \sigma_{\hat{f}_{h_i}}(\mathbf{x}, h_i).$$
2. For $i = 1, 2, \dots, J - 1$, let

$$\bar{L}_{i+1} = \max\{\bar{L}_i, L_{i+1}\}, \quad \bar{L}_1 = L_1,$$

$$\underline{U}_{i+1} = \min\{\underline{U}_i, U_{i+1}\}, \quad \underline{U}_1 = U_1.$$
3. The ICI rule is finding the largest i , when $\bar{L}_i \leq \underline{U}_i$, $i = 1, 2, \dots, J$, is still satisfied.

4. PROPOSED METHOD

We propose applying the ICI approach to optimize the spatial adaptation for local region-based active contours. For each point, the ICI approach finds an optimal kernel size that permits the trade-off between the bias and the variance of estimates. This optimal local scale is then used for the estimation of the local means and variance of the segmentation model as given in Eq. (2). In this context, window size and scale are interchangeable concepts.

Suppose a noisy image I with intensity inhomogeneities in its foreground and background. We define an initial zero level set C , as the yellow contour shown in Fig. 1. Given a finite set of window sizes \mathbf{h} , we calculate the g_h for each element. Then utilize Eq. (3) to get the local estimations of the regions Ω_i and Ω_o respectively. It means that if a point \mathbf{x} is inside of C , its approximation uses the local interior $\Omega_i^l(\mathbf{x}) = K_h(\mathbf{x}) \cap \Omega_i$, and vice versa. As introduced in sec. 3, we can calculate the confidence intervals of this estimation, then apply the ICI algorithm for each point. After that, we obtain the optimal kernel size that well balances the bias-variance trade-off.

To find out the relation of these data adaptive scales with the position of the segmentation contour, we picked out several typical points for analysis. As we are only interested on a narrow band of C , within which we select 4 pairs of neighbors where one is in Ω_i (green '+') and the other is in Ω_o (red 'o'). The corresponding estimated scales are also illustrated on the same figure with circles (Fig. 1, left). The leftmost pair P_4 lay around a region with very low contrast between Ω_i and Ω_o , where the local statistics are very similar. Also the contour near P_4 is the correct boundary, where the segmentation is locally achieved. In order to maintain this partition, we tend to consider more information, which is corresponding to

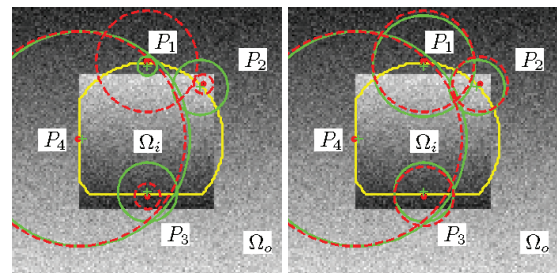


Fig. 1. Adaptive local kernel sizes. Left: obtained by LPA-ICI. Right: used in the segmentation.

the larger kernel size obtained by LPA-ICI algorithm. For the pair P_1 around the top of C , the inside one laying between C and the real boundary has small kernel size. This is expected, as in that position, larger windows will introduce greater estimation bias. Its symmetric point has, however, larger scale estimates, because in its neighboring region within Ω_o , the image is relatively homogenous. The pairs P_2 and P_3 have, however, larger window size inside and smaller one outside.

Therefore, if we directly use these kernel sizes in the segmentation algorithm, as C gets closer to the real boundary, the local regions of points between them are decreasing, and so will be the estimated local scale. This brings out the problem that the closer C is to the correct segmentation, the slower the evolution speed is. Analyzing case P_1 (P_2), the scale inside (outside) has to be at least as big as the outside (inside) in order to increase the force driving the segmentation process. To accelerate the evolution, we run a max filter of a small size 3×3 , on the estimated local scales, so that near C , the used scales have similar values (Fig. 1, right). This filtering operation is necessary only when the algorithm is in progress. Indeed, the estimated scales are very appropriate as it can be seen on the correct segmentation P_4 .

Our algorithm is organized as follows:

1. **Initialization:** Given an image I , an initial segmentation C or ϕ , a finite set \mathbf{h} and a vector of polynomials ψ . For each $h \in \mathbf{h}$ calculate the LPA kernel g_h .
2. **Optimal spatial kernel size estimation–LPA-ICI:** Estimate respectively the image inside and outside of C with g_h . Apply the ICI rule in order to get \hat{h} .
3. **Local region-based segmentation:** Use the local statistics within \hat{h} for full MS segmentation [6].
4. Repeat steps 2 and 3 until convergence.

5. EXPERIMENT AND DISCUSSION

This section demonstrates the performance of the proposed adaptive segmentation algorithm and compare it with the single local scale method [6], which explains the importance of the adaptive selection of local scales in order to achieve acceptable results. We denote by \hat{h} the half window size and by \hbar the kernel standard deviation.

Fig. 2 shows a desirable segmentation result of the image in Fig. 1 with the proposed method. For a further analysis, plots of the estimated local window sizes \hat{h} along the final contour, for the inside and the outside regions are also shown. These plots follow a pixel clockwise parametrization of the curve starting from the top right corner. Notice for example that the estimated scales in Ω_i are smaller than the outside region. This can be explained by the fact that the inhomogeneity in Ω_i is stronger than that in Ω_o . This difference of scales, between the inside and the outside, is important for the forces in competition around low contrasted boundaries.

Then, we chose to test the proposed method on US B-mode image, which are known to have low signal-to-noise ratio, low contrast, and high amounts of speckle. Consequently,

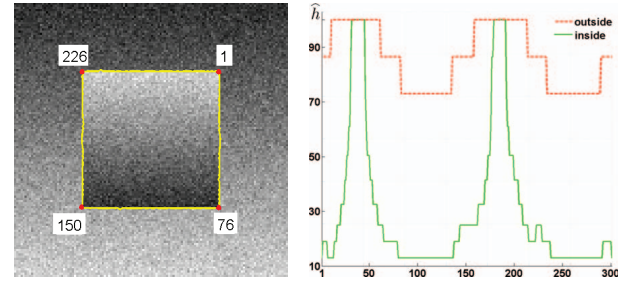


Fig. 2. Left: segmentation result of the proposed method, the initial curve is shown in Fig. 1. Right: the estimated optimal spatial kernel sizes, clockwise along C , for the Ω_i and Ω_o .

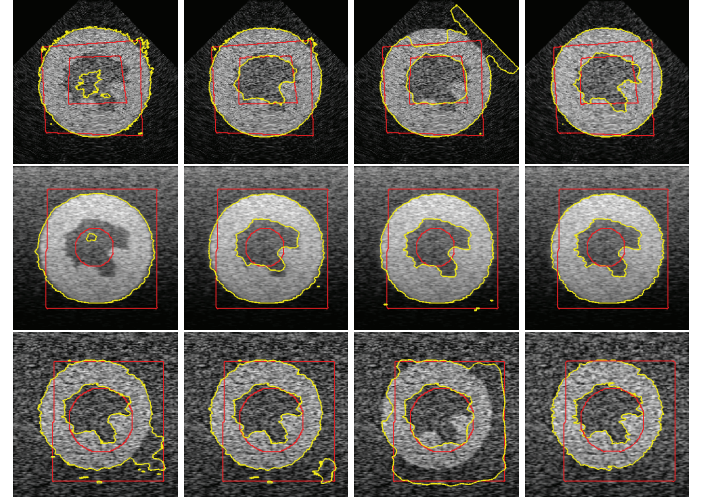


Fig. 3. Segmentation of simulated US images. From left to right: local MS with the $\hbar = 70, 40, 10$; the proposed method. From top to bottom: ‘PH1_45dB’, ‘PH2_ATT2.65dB’ and ‘PH3_35dB’ with 3 different initializations.

US image segmentation is strongly influenced by the quality of data [13]. Our main objective, is not to solve this difficult problem, but to show that adaptive scale selection gives an acceptable solution to the attenuation problem. Therefore, we chose to demonstrate the proposed approach on realistic US simulations. We have used the US simulation program Field-II [14] to synthesize data with known ground truth. For example, the first phantom (PH1) of size 100 mm cube was placed at 10 mm depth from the transducer surface, and was scanned with a 7 MHz 128 elements phased array transducer. A single transmit focus at 60 mm from the transducer was used, and focusing during reception is at 10 to 150 mm in 1 mm increments. The images consist of 128 lines with 0:7 degrees between lines. Hanning apodization in transmit and receive was used in all experiments. Three levels of tissue attenuations were simulated for each phantom. We also used different dB ranges for the envelope log compression to simulate different image contrasts.

Fig. 3 shows some segmentation results obtained by the local MS model for 3 scales and by the proposed adaptive scale algorithm. Notice the influence of the choice of size of the windows for the local method with single scale. Our proposed scale selection gives better results.

In order to further study the influence of the kernel size

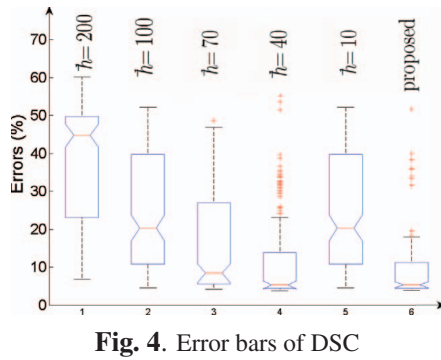


Fig. 4. Error bars of DSC

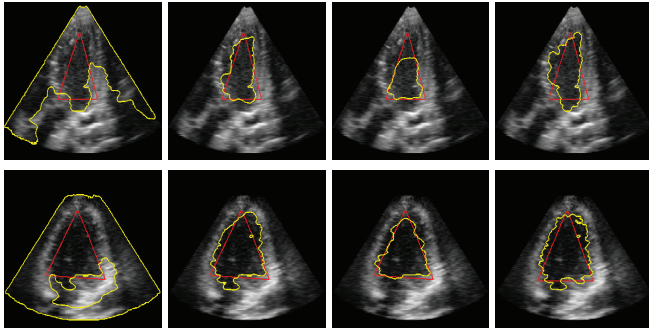


Fig. 5. Segmentation of US images. Top line from left to right: local MS method with $h = 6, 4, 2$ and the proposed method; bottom one: local MS method with $h = 9, 8, 3$ and the proposed method.

in single scale local region-based method, Fig. 4 shows the quantitative comparison between the proposed approach to the single scale LMS method. The error bar shows statistics of the Dice similarity Coefficient ($1 - \text{DSC} = 1 - 2 \frac{|S \cap S_{ideal}|}{|S| + |S_{ideal}|}$) of 60 images of the simulated 3 phantoms with 3 different initializations (overall $60 \times 3 = 180$ images). The DSC statistics shows clearly that, when $h = 200, 100$, the local MS model is equivalent to the global Gaussian model in [2]. For the scale $h = 70, 10$, as the images shown in Fig. 3, the segmentations have high variability of errors. For these test images, the local MS method using $h = 40$ provides a significant improvement with respect to the other 4 kernel sizes, however, its DSC errors still have larger variance and more outliers than the proposed adaptive one. This is expected since the appropriate scale should depend on the specific cases, such as the position of the initial C , the noise level and the size of the target. Therefore, the proposed adaptive segmentation algorithm has an increase of robustness to initialization and to attenuation.

Finally, we show some preliminary results on real US images in Fig. 5. The local MS method, using a fixed scales statistics, has varying results. An appropriate scale is necessary in the application of local region-based segmentation, while the proposed method need not.

6. CONCLUSIONS

In this paper, we presented a local region-based segmentation method with adaptive scale selection within the level set framework. The ICI rule is used to derive the optimal scales for interior and exterior of the current segmentation. We showed the results of the proposed method with comparisons

to the single scale local MS model, and the statistics of the segmentations for 60 simulated US images with 3 different initializations. All of these illustrate the improvement on the state of the art of segmentation methods. Although our experiments on the real US images are limited, it suggest the proposed method has potential application. The continued development of the proposed method, with more experiments, should be researched in the foreseeable future.

Acknowledgements: The authors would like to thank Professor Alison Noble from the BioMedIA Laboratory, University of Oxford, for providing us the real US data. Financial supports of this research from the Chinese Scholarship Council and the Laboratoire Heudiasyc are greatly appreciated.

7. REFERENCES

- [1] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, 2001.
- [2] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," Tech. Rep. 4695, INRIA, 2003.
- [3] A. Yezzi, A. Tsai, and A. Willsky, "A fully global approach to image segmentation via coupled curve evolution equations," *J. Visual Communication Image Representation*, vol. 13, no. 1-2, pp. 195–216, 2002.
- [4] C. Li, C. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, 2008.
- [5] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2029–2039, 2008.
- [6] T. Brox and D. Cremers, "On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional," *Int. J. Comput. Vis.*, vol. 84, no. 2, pp. 184–193, 2009.
- [7] L. Wang, C. Li, Q. Sun, D. Xia, and C. Kao, "Active contours driven by local and global intensity fitting energy with application to brain MR image segmentation," *Comput. Med. Imaging Graph.*, vol. 33, no. 7, pp. 520–531, 2009.
- [8] D. Boukerroui, O. Basset, A. Baskurt, and A. Noble, "Segmentation of echocardiographic data. Multiresolution 2D and 3D algorithm based on gray level statistics," in *MICCAI*, 1999.
- [9] D. Boukerroui, A. Baskurt, A. Noble, and O. Basset, "Segmentation of ultrasound images—multiresolution 2D and 3D algorithm based on global and local statistics," *Pattern Recognit. Lett.*, vol. 24, pp. 779–790, 2003.
- [10] Q. Yang and D. Boukerroui, "Optimal spatial adaptation for local region-based active contours: An intersection of confidence intervals approach," in *VISIGRAPP*, 2011, pp. 87–93.
- [11] D. Mumford and J. Shan, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, pp. 577–685, 1989.
- [12] V. Katkovnik, K. Egiazarian, and J. Astola, "Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule," *J. Math. Imag. Vis.*, vol. 16, no. 3, pp. 223–235, 2002.
- [13] A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imaging*, vol. 25, no. 8, pp. 987–1010, 2006.
- [14] J. A. Jensen, "Field: A program for simulating ultrasound systems," in *10th Nordic-Baltic Conf. on Biomedical Imag.*, 1996, vol. 34, pp. 351–353.

A local Rayleigh model with spatial scale selection for ultrasound image segmentation

A local Rayleigh model with spatial scale selection for ultrasound image segmentation

Djamal Boukerroui
<http://www.hds.utc.fr/~dboukerr>

Université de Technologie de
Compiègne
Heudiasyc UMR CNRS 7253
BP 20529 - 60205 Compiègne Cedex,
France.

Abstract

Ultrasound images are very noisy, with poor contrast and the attenuation of the acoustic wave in the depth of the observed medium leads to strong inhomogeneities in the image. Segmentation methods using global image statistics give unsatisfactory results. The use of local image statistics can solve effectively the problem of attenuation. The contribution of this paper is two folds. First, we propose the study of the adaptation of the global model proposed by Sarti et al. [28]. We kept the variational framework and the Rayleigh model of the observed image statistics. Second, we propose an interesting and generic adaptive scale selection algorithm based on the Intersection of Confidence Interval rule. The latter is also applied to the local Gaussian segmentation model of Brox and Cremers. Results on realistic simulations of ultrasound images show the robustness and the superiority of the local Rayleigh model. The efficiency and the genericity of the proposed scale selection strategy is also demonstrated.

1 Introduction

In medical imaging, the ultrasonic wave is modelled as a progressive plane mechanical wave. When this wave encounters an interface between two tissues with different acoustic characteristics, a part of the incident wave is reflected (specular echoes) in the direction of the probe. Along with these specular echoes, backscattered echoes are added by the microscopic structure of the medium. Backscattering is the origin of the speckle phenomenon, which characterizes ultrasound imaging with a granular appearance. The speckle is a multiplicative noise, strongly correlated and more importantly, with non Gaussian statistics. These characteristics differ greatly from the traditional assumption of white additive Gaussian noise, often taken in ultrasound image segmentation, which leads to reduction of the effectiveness of the methods. Thus, several researchers have studied the statistics of the envelope image of the received echo signal in order to obtain processing algorithms specific to ultrasonic data [24, 25]. Several distribution families have been proposed in the literature: specific models such as the Rayleigh [6, 32], Rice [14, 32], Nakagami [30], K-distribution [15, 29, 30] and more general models, the Homodyne K-distribution [9] and the more recently RIIG [10, 11]. Notice the large variability of the proposed models, which is due to the strong dependence of the observed statistics on the density of scatters and on their spatial distribution (uniform

2 *BOUKERROUI: A LOCAL RAYLEIGH MODEL WITH SPATIAL SCALE SELECTIONS*

or random) in the analysed tissue [32]. All these models are for the envelope of the received echo signal (i.e before interpolation, log-compression and Time-Gain-Compensation). Thus, the validity of such models on ultrasound images acquired under clinical conditions is questionable [22, 31, 36].

Image segmentation in general is an ill-posed problem in the sense of Hadamard and difficult to solve. This problem is especially difficult when the data to be processed are medical in nature. The literature on the subject is abundant and the performances of the proposed techniques are highly dependent on the quality of data [25, 34]. Ultrasound data are very noisy, with poor contrast, and often presents missing boundaries of the object of interest due to problems of specular reflection, shadows, signal dropout and attenuation. As a consequence, conventional intensity gradient-based methods have had limited success on typical clinical images [25]. Solutions using the phase information, theoretically invariant to image contrast, are successfully used in [2, 23]. Note also that segmentation methods based on global statistical models, regardless of the used framework, fail on this type of data, mainly because of the attenuation problem. Adaptive solutions robust to attenuation exist in the literature [1, 3, 4, 25]. Local image statistics were used for the estimation of the segmentation model's parameters.

Recently, there has been a reinvestigation of the use of local statistics by the image segmentation community, but in a variational framework [5, 19, 21, 27]. For instance, the adaptation of the binary model of Chan and Vese [8] is presented in [21]. A comprehensive formalization of generic local segmentation models is proposed in [19], with examples of local energies derived from global ones. Also, a statistical interpretation of the piecewise smooth segmentation model of Mumford-Shah using local Gaussian models has been proposed in [5]. An interesting alternative with local Gaussian distributions is also proposed in [33]. These recent studies show a better behavior of these local models on images with strong intensity inhomogeneities. This contribution falls under this context. We propose the study of adaptation of the model proposed by Sarti et al. [28]. The latter assumes a global Rayleigh model envelope image statistics. The reader is also referred to [20] for a generalisation to the exponential family.

2 Segmentation model

Let $I : \Omega \rightarrow \mathbb{R}^+$ denote a given observed image and \mathcal{C} be a closed contour represented as the zero level set of a signed distance function ϕ , i.e., $\mathcal{C} = \{\mathbf{x} | \phi(\mathbf{x}) = 0, \mathbf{x} \in \Omega\}$. The interior Ω_i and the exterior Ω_e of \mathcal{C} are defined by a smooth approximation of the Heaviside function respectively by: $H_i(\phi) = H(\phi)$ and $H_o(\phi) = 1 - H(\phi)$. Image intensities are supposed to be independent realisation of random variables with a certain probability density function (pdf) $p(I)$. We seek the partition of Ω that maximizes the likelihood function of the observed data. Given the independence assumption, this leads to the minimization of the following energy function [28]

$$E(\phi) = - \sum_{r \in \{i,o\}} \int_{\Omega} H_r(\phi) \log p(I(\mathbf{x})) d\mathbf{x} + \lambda \int_{\Omega} \delta(\phi) |\nabla \phi| d\mathbf{x}, \quad (1)$$

where the first two terms are the data terms and the last one is a length regularisation term, added in order to keep the curve smooth, with a positive weight penalty λ . We will further

assume that the random intensity $I(\mathbf{x})$ follows a Rayleigh pdf with a parameter σ^2 :

$$p(I(\mathbf{x})) = \frac{I(\mathbf{x})}{\sigma^2} \exp\left(-\frac{I(\mathbf{x})^2}{2\sigma^2}\right) \quad \text{and} \quad \widehat{\sigma^2}_{\text{ML}} = \frac{\int_{\Omega_r} I(\mathbf{x})^2 d\mathbf{x}}{2 \int_{\Omega_r} d\mathbf{x}}, \quad (2)$$

where $\widehat{\sigma^2}_{\text{ML}}$ is a Maximum Likelihood estimates under the assumption that all the observed pixels in the domain Ω_r are identically distributed. In the work of Sarti et al. [28], only two global domains were used, Ω_i for the inside and Ω_e for the outside pixels. Therefore the hypothesis of identically distributed observations is generally false for ultrasound images because of the presence of strong intensity inhomogeneities due to attenuation and other factors. However, the assumption remains true if the estimate is made locally in a region centered around each pixel of the domain Ω . Thus the energy corresponding to the inside term of (1) is given by:

$$E_i(\phi) = \int_{\Omega} H(\phi) \left[\frac{I(\mathbf{x})^2}{2\sigma_i^2(\mathbf{x})} + \log(\sigma_i^2(\mathbf{x})) \right] d\mathbf{x} \quad (3)$$

$$\text{and} \quad \sigma_i^2(\mathbf{x}) = \frac{\int_{\Omega} H(\phi) K(\mathbf{x} - \xi) I(\xi)^2 d\xi}{2 \int_{\Omega} H(\phi) K(\mathbf{x} - \xi) d\xi}, \quad (4)$$

here $K(\cdot)$ is any given kernel defining the spatial locality around the position \mathbf{x} . In this work a Gaussian kernel with a standard deviation σ_K is used. For computing the Euler-Lagrange equation of (3), we employ the Gâteaux derivative. Following [5], for any point \mathbf{x} and for any perturbation $\psi(\mathbf{x})$ we have:

$$\begin{aligned} \frac{\partial E_i(\phi + \varepsilon \psi)}{\partial \varepsilon} \Big|_{\varepsilon \rightarrow 0} &= \int_{\Omega} \delta(\phi(\mathbf{x})) \left[\frac{I(\mathbf{x})^2}{2\sigma_i^2(\mathbf{x})} + \log(\sigma_i^2(\mathbf{x})) \right] \psi(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\Omega} H(\phi(\mathbf{x})) \left[\frac{2\sigma_i^2(\mathbf{x}) - I(\mathbf{x})^2}{2\sigma^4(\mathbf{x})} \right] \sigma_\phi^2(\mathbf{x}) d\mathbf{x}, \end{aligned} \quad (5)$$

$$\text{where} \quad \sigma_\phi^2(\mathbf{x}) = \frac{\int_{\Omega} \delta(\phi(\xi)) K(\mathbf{x} - \xi) (I(\xi)^2 - \sigma_i^2(\mathbf{x})) \psi(\xi) d\xi}{2 \int_{\Omega} H(\phi(\mathbf{z})) K(\mathbf{x} - \mathbf{z}) d\mathbf{z}}. \quad (6)$$

Using the following notations:

$$\begin{aligned} F_1(\mathbf{x}) &= \int_{\Omega} K(\mathbf{x} - \mathbf{y}) H(\phi(\mathbf{y})) d\mathbf{y} = (K * H(\phi))(\mathbf{x}), \\ F_2(\mathbf{x}) &= \left(\bar{K} * \frac{H(\phi)(2\sigma^2 - I^2)}{\sigma^4 F_1} \right)(\mathbf{x}), \quad F_3(\mathbf{x}) = \left(\bar{K} * \frac{H(\phi)(2\sigma^2 - I^2)}{\sigma^2 F_1} \right)(\mathbf{x}), \end{aligned}$$

where \bar{K} is the mirror version of K and $*$ is the convolution operation. The Euler-Lagrange equation for the minimization of the energy E_i , given in (3), obtained from (5) simplifies then as follows:

$$0 = \delta(\phi) \left[\frac{I(\mathbf{x})^2}{2\sigma_i^2(\mathbf{x})} + \log(\sigma_i^2(\mathbf{x})) + \frac{1}{4} (I(\mathbf{x})^2 F_2(\mathbf{x}) - F_3(\mathbf{x})) \right],$$

and can be implemented very efficiently using recursive gaussian filtering [12]. The derivation of the Euler-Lagrange equation corresponding to the energy in Ω_o (second term in the sum in (1)) is obtained by analogy to the above equation. The gradient of the length term can be found in [8, 28].

3 Spatial scale selection

Local region-based segmentation models are surely a better alternative to global ones in the presence of intensity inhomogeneities. Such models however may be more sensitive to initialisation if the chosen local spatial scale is not appropriate. A decrease of robustness to noise is also observed when small scales are used. To our knowledge, two pixel dependent scale selection methods have been introduced recently. Piovano and Papadopoulos defined the local scale as “the smallest one inducing an evolution speed superior to a given threshold” [27]. Yang and Boukerroui proposed that the optimal scale is “in the sense of the mean-square error minimisation of a Local Polynomials Approximation (LPA) of the observed image conditional on the current segmentation” [35]. The former is a generic approach and the latter is based on an Intersection of Confidence Intervals (ICI) algorithm, which finds the optimal LPA estimate of the image. Since, the LPA assumes an additive gaussian noise model, the performances of this method may decrease when the noise model is not additive gaussian. In this section, we first recall Piovano and Papadopoulos’s methods and then introduce a new generic ICI rule for the spatial scale selection. Both methods can estimate continuous scales but for efficiency reasons, the search is generally limited to a finite set of ordered scale values $\mathbf{h} = \{h_1 < h_2 < \dots < h_J\}$.

3.1 Evolution speed threshold based approach

In [27], the idea is to find the most salient scale to make the contour locally evolve. It is defined as the smallest one inducing an evolution speed superior to a given threshold ε :

$$h^*(\mathbf{x}) = \inf_{h \in \mathbf{h}} \{h : |\partial_t \phi(\mathbf{x})| > \varepsilon \quad \text{and} \quad \sigma_K = h\} . \quad (7)$$

Thus, at each iteration of the level set evolution, the optimal scale is found for every image pixel along the zero level set function, by increasing $\sigma_K(\mathbf{x})$ from the minimum h_1 to the maximum h_J , until the absolute value of the evolution speed is superior to ε . It is unclear how to fix the positive constant ε .

3.2 ICI rule

In the following section, we introduce a new alternative approach in order to choose a good value for the spatial scale at every pixel location along the contour. Our approach is based on the idea of choosing the largest scale that gives the best estimate of the segmentation model parameters. The approach is generic and can be applied to any parametric pdf. For the sake of clarity, we detail the development for the Rayleigh case only and we give the main results for the Gaussian case.

3.2.1 The Rayleigh case

It can be shown that the ML estimator of σ^2 given in eq. (2) is *efficient*. It is therefore unbiased and has the lowest possible variance defined by the Cramer-Rao bound. It is also asymptotically Gaussian as it is a ML estimator: for $n \rightarrow \infty$ $\widehat{\sigma^2} \rightarrow \mathcal{N}(\sigma^2, \sigma^4/n)$ where $n = \int_{\Omega_r} d\mathbf{x}$ in our context. Therefore we can estimate a Confidence Interval on the estimates

of σ^2 by :

$$P \left(\sigma^2 \in \left[\frac{\widehat{\sigma^2}}{1 + \frac{u_{1-\alpha/2}}{\sqrt{n}}}, \frac{\widehat{\sigma^2}}{1 - \frac{u_{1-\alpha/2}}{\sqrt{n}}} \right] \right) \simeq 1 - \alpha . \quad (8)$$

where u_α is the α^{th} percentile of the standard Gaussian distribution $\mathcal{N}(0, 1)$. Equivalently, the confidence interval for the estimator $\sigma_i^2(\mathbf{x})$ given in eq. (4) is given by eq. (8) with

$$\widehat{\sigma^2} = \sigma_i^2(\mathbf{x}) \quad \text{and} \quad n = \frac{(\int_{\Omega} H(\phi) K(\mathbf{x} - \xi) d\xi)^2}{\int_{\Omega} (H(\phi) K(\mathbf{x} - \xi))^2 d\xi} .$$

Therefore in an ideal situation, where the observed data is identically distributed, the bigger n is, the better is the estimates of σ^2 . Bear in mind that the hypothesis of identically distributed data in the local window will become less and less valid as the scale of the kernel K grows and will lead to an increasingly biased estimations. This means that there exists a bias-variance balance that gives the ideal scale. We can make use of the ICI algorithm to search for the largest local window (minimising variance) that gives us the best estimate of σ^2 (minimising bias). First, the estimates $\widehat{\sigma^2}_h$ and their corresponding CIs are calculated for all $h \in \mathbf{h}$. Then, the ICI rule identifies the best scale h^* as the largest in \mathbf{h} that has a non empty intersection of CIs with all scales in $\{h_i \in \mathbf{h}, h_i < h^*\}$. The ICI algorithm is defined by the following steps [13, 18]:

1. Define a sequence of confidence intervals $Q_i = Q(h_i)$ with their lower L_i and upper bounds U_i using eq. 8.
2. For $i = 1, 2, \dots, J - 1$, let

$$\begin{aligned} \bar{L}_{i+1} &= \max\{\bar{L}_i, L_{i+1}\}, & \bar{L}_1 &= L_1 , \\ \underline{U}_{i+1} &= \min\{\underline{U}_i, U_{i+1}\}, & \underline{U}_1 &= U_1 . \end{aligned}$$

According to these formulas, \bar{L}_{i+1} and \underline{U}_{i+1} are respectively non-decreasing and non-increasing sequences.

3. The ICI rule is finding the largest i , when $\bar{L}_i \leq \underline{U}_i$, $i = 1, 2, \dots, J$, is still satisfied.

3.2.2 The Gaussian case

When the observed image intensities $I(\mathbf{x})$ are supposed to follow local Gaussian distributions [5], we can base our scale estimation on the confidence intervals of the local means. Recall that if $X \sim \mathcal{N}(\mu, \sigma^2)$ and given an independent and identically distributed sample of size n then:

$$P \left(\mu \in \left[\bar{X} - \frac{S}{\sqrt{n}} t_{n-1, 1-\alpha/2} , \bar{X} + \frac{S}{\sqrt{n}} t_{n-1, 1-\alpha/2} \right] \right) = 1 - \alpha . \quad (9)$$

where S^2 is the unbiased estimate of the variance and $t_{n,\alpha}$ is the α^{th} percentile of a Student's t-distribution of n degrees of freedom [26]. Thus, Eq. 9 gives us a set of CIs that can be utilised in an ICI algorithm in order to select the best scale for the estimation of the local means. Note that eq. 9 is exact $\forall n$. By use of the Central Limit Theorem, eq. 9 can be used as an approximation of the CIs by replacing Student's percentiles by the Gaussian one in the general non gaussian case.

3.2.3 The scale selection procedure

The above proposed ICI rule is used to estimate the best spatial kernel at every pixel location along the contour. Note however that the ICI rule will produce two scale values, h_i^* and h_o^* corresponding for the inside region Ω_i and the outside region Ω_o respectively. In [35] the used scale corresponds to the maximum between the two values. In this work we followed the same procedure as in [35] until converges, we ran few iterations by replacing the maximum by an average filtering and then by using directly the estimated scales: h_i^* for the inside energy term and h_o^* for the outside energy term.

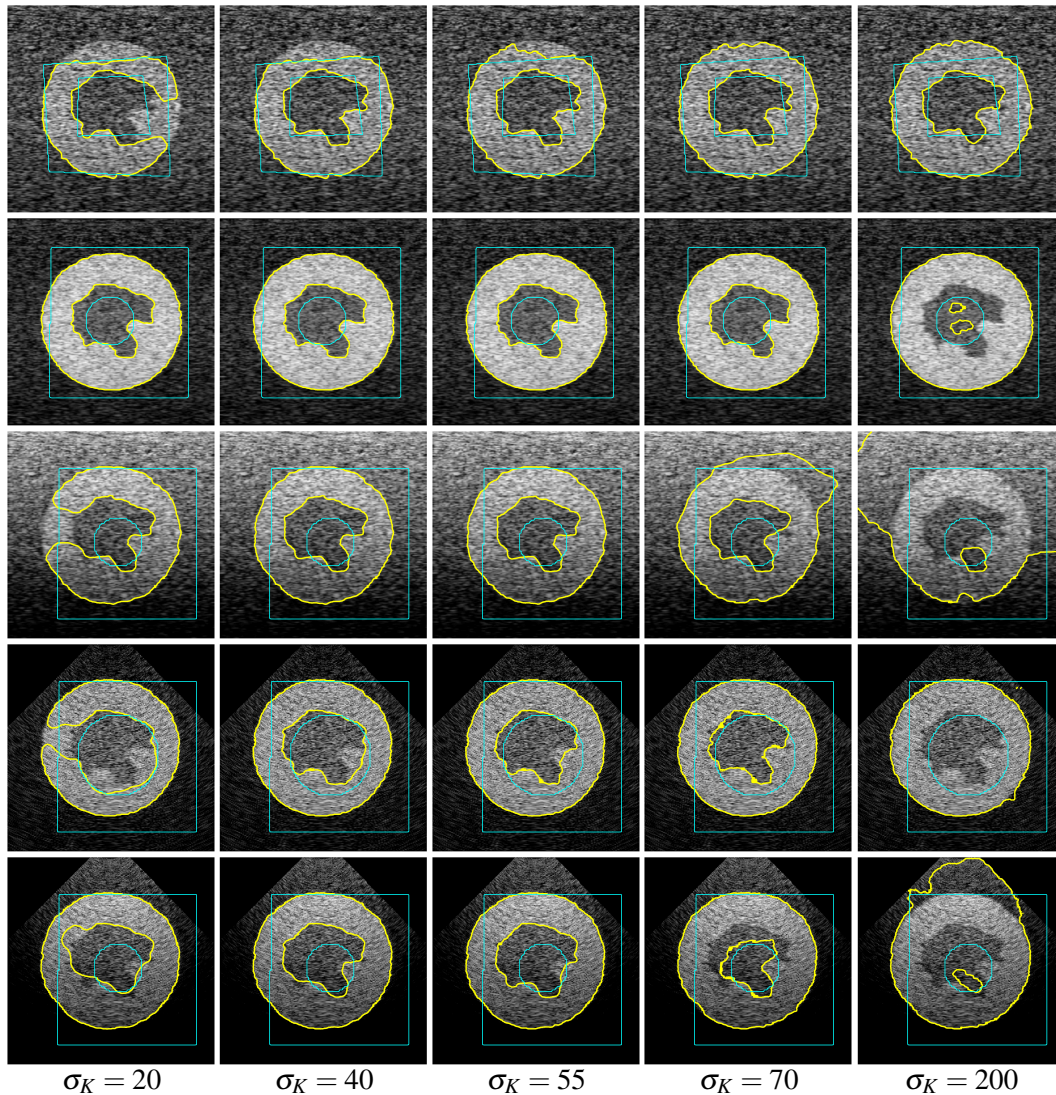


Figure 1: Illustrative segmentation results of the local Rayleigh model on simulated US images with different: tissues characteristics, attenuation level and log compression parameters. Thin lines show the (4 different) initialisations. For better contrast, the images are displayed after the logarithmic compression of the envelope but the image segmentation algorithm uses the data before compression.

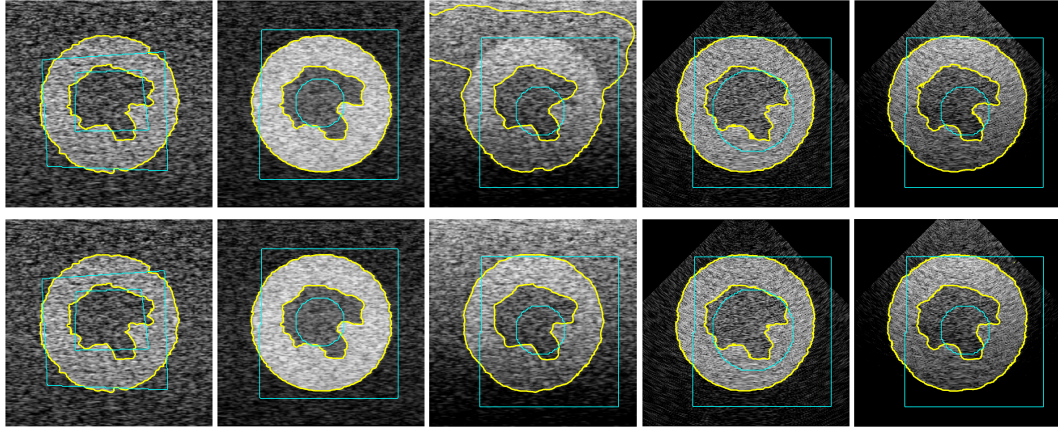


Figure 2: Illustrative segmentation results of the proposed local Rayleigh model with spatial scale selection. First line: scale selection method of [27] with $\varepsilon = 0.5$. Second line: the proposed ICI based approach with $\alpha = 0.05$. These results are to be compared to the one shown in Fig. 1. The used set $\mathbf{h} = \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80\}$ for an image size of 256×256 .

4 Results

In order to demonstrate the usefulness of the proposed approach and quantify its performances, we chose to test it on realistic US simulations. To this end, we have used the simulation program Field-II [16, 17], to synthesize phantom data with known ground truth. A linear scan of a first phantom (PH1) was done with a 290 elements transducer using 64 active elements. The scatterers in the phantom were randomly distributed within the phantom of $80 \times 80 \times 15$ mm cube size. 128 lines were simulated at 5 Mhz. The second phantom (PH2) of size $100 \times 100 \times 15$ mm cube was placed at 10 mm depth from the transducer surface, and was scanned with a 7 MHz 128 elements phased array transducer. The images consist of 128 lines with 0.7 degrees between lines. Hanning apodization in transmit and receive was used in all experiments. Two scatterers amplitudes with three levels of tissue attenuations were simulated for both phantoms. We also used several dB ranges for the envelope logarithmic compression to simulate different image contrasts. Some typical images are shown in Fig. 1.

The first experiment demonstrates the behaviour of the proposed local Rayleigh approach when different sizes of the local spatial kernel and different initialisations were used (see Fig. 1). Small local scales generally decrease the capture range of the active contour and also increase the number of local minima. Big scale values lead to unsatisfactory results in the presence of intensity inhomogeneities. Figure 2 shows the corresponding results to the one shown in Fig. 1, when an automatic scale selection procedure is used. Observe that the proposed ICI rule based approach outperform the method in [27].

In order to understand better the behaviour of the two automatic scale selection algorithms, Fig. 3 shows the contour evolution and the estimated scale maps for the two approaches on the image of Fig. 2(middle column). Although, the interpretation of the scale maps estimated by the proposed ICI rule is possible, it is difficult to understand the ones obtained with Piovano and Papadopoulos's method. Indeed, our approach is driven by a bias-variance tradeoff for the local estimation of the segmentation model parameter, conditional on the current segmentation map. This is why for example we get higher scales when the algorithm converges (except the last iteration where we deliberately reduce the scales in order

8 *BOUKERROUI: A LOCAL RAYLEIGH MODEL WITH SPATIAL SCALE SELECTIONS*

to increase accuracy, see sec.3.2.3). Furthermore, it is easier to set the value of α , because it is more interpretable, then the threshold ε . For this example, the segmentation fail, with Piovano and Papadopoulos's scale selection method, because the estimated scales are too big in the upper part of the image.

The quantitative evaluation of the proposed approach compared to ground truth are summarized in Table 1. It shows statistics of the Dice similarity Coefficient $DSC(S, S_{ideal}) = 2 \frac{|S \cap S_{ideal}|}{|S| + |S_{ideal}|}$ and the Mean Absolute Distance (MAD) [7] of segmentation results of the 60 simulated images of the two phantoms with 4 different initialisations ($60 \times 4 = 240$ results). The closer the DSC and the MAD values to 1 and 0 respectively, the better is the segmentation. The table shows the minimum, the three Quartiles, the maximum and the interquartile range (IQR) of the measurements, for six sizes of the spatial kernel $K(\cdot)$. This table clearly shows that local image statistics should be used in the presence of intensity inhomogeneities. The minimum and the Q_1 values of DSC and the maximum and the Q_3 values of MAD (indicating the worst cases) are very informative. Notice also the IQR values, a robust measure of dispersion. The worst results are for very large values of σ_K . In this case the algorithm's behaviour is similar to the global method of Sarti et al. [28]. Notice that both scale selection methods performed pretty well. These measurements indicates a superiority of the new proposed scale section methodology on Piovano and Papadopoulos's approach. The interpretation of these data suggests the use of local statistics for the segmentation of ultrasound images.

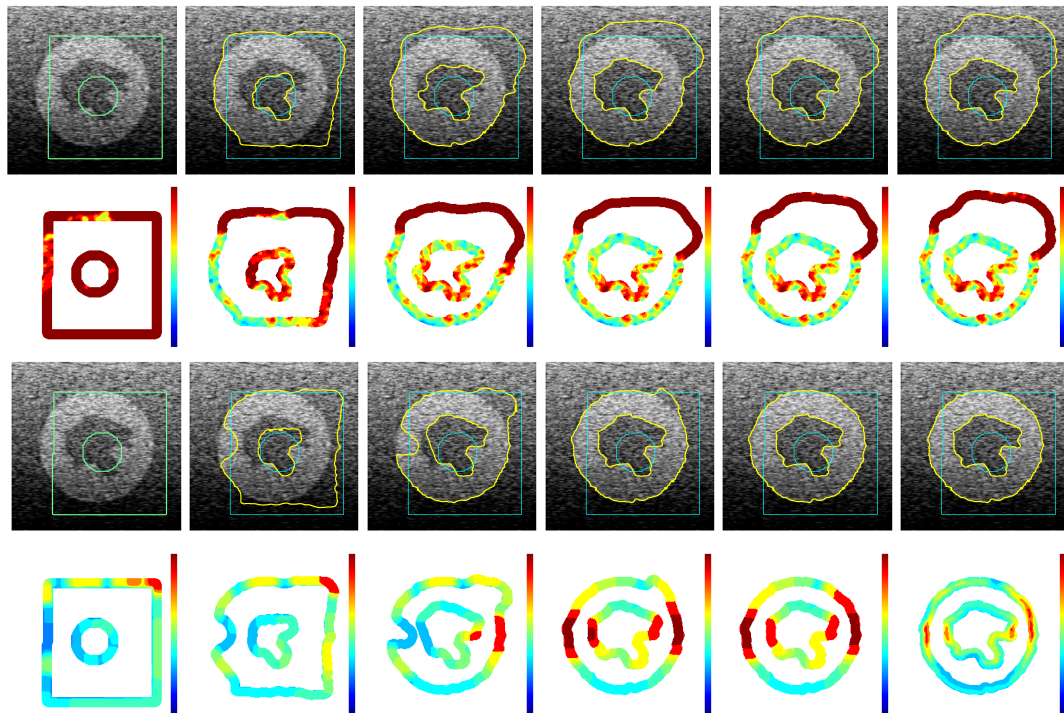


Figure 3: Behaviour of the two automatic scale selection algorithms. 1st and 3rd lines: contour evolution of the local Rayleigh segmentation model for iterations 1, 10, 25, 50, 75 and 120, respectively when using the scales shown in the 2nd (estimated with Piovano and Papadopoulos's method, see sec.3.1) and the 4th (obtained with the proposed ICI rule) lines.

	DSC×100						MAD (in pixels)					
	min	Q ₁	Q ₂	Q ₃	max	IQR	min	Q ₁	Q ₂	Q ₃	max	IQR
$\sigma_K = 20$	56.4	83.5	89.2	93.5	96.5	10.0	1.10	1.33	2.09	3.87	20.51	2.54
$\sigma_K = 30$	60.2	90.3	95.8	96.0	96.5	5.8	1.11	1.29	1.38	1.95	23.70	0.66
$\sigma_K = 40$	46.3	95.4	95.9	96.1	96.5	0.7	1.11	1.25	1.38	1.48	23.39	0.23
$\sigma_K = 55$	54.2	95.4	95.7	96.0	96.5	0.6	1.10	1.29	1.39	1.53	24.76	0.23
$\sigma_K = 70$	62.7	93.7	95.1	95.7	96.4	2.0	1.06	1.37	1.53	1.74	24.19	0.37
$\sigma_K = 200$	46.1	54.5	79.3	85.6	95.1	31.1	1.29	2.26	4.06	19.76	65.39	17.50
$\sigma_K = h^*$	62.4	93.6	95.9	96.2	96.5	2.6	1.11	1.21	1.31	1.90	24.27	0.69
$\sigma_K = h_{ICI}^*$	68.2	95.5	96.1	96.2	96.5	0.7	1.09	1.29	1.36	1.47	21.54	0.17

Table 1: Statistics of the DSC and the MAD measures obtained on 60 images with 4 different initialisations. σ_K is the standard deviation of the spatial gaussian kernel. h^* is the proposed segmentation approach with the scale selection in [27]. h_{ICI}^* is the proposed approach with the proposed ICI scale selection procedure. $\sigma_K = 200$ is equivalent to the method in Sarti et al. [28].

	DSC×100						MAD (in pixels)					
	min	Q ₁	Q ₂	Q ₃	max	IQR	min	Q ₁	Q ₂	Q ₃	max	IQR
$\sigma_K = 10$	62.5	72.0	90.5	94.6	96.1	22.7	1.11	3.86	15.06	17.93	34.50	14.07
$\sigma_K = 20$	65.9	88.4	94.8	95.8	96.3	7.5	1.14	1.61	7.78	19.80	30.33	18.19
$\sigma_K = 30$	49.4	93.9	95.6	95.9	96.5	1.9	1.12	1.43	1.94	12.61	27.49	11.18
$\sigma_K = 40$	58.8	89.8	94.8	95.7	96.4	5.9	1.15	1.48	2.02	10.78	25.20	9.31
$\sigma_K = 55$	54.0	80.9	93.3	95.4	96.0	14.5	1.26	1.51	2.02	10.08	27.43	8.56
$\sigma_K = 70$	50.3	72.3	91.7	95.1	95.8	22.8	1.26	1.64	2.39	15.10	51.42	13.46
$\sigma_K = 200$	42.0	50.6	64.5	80.9	95.0	30.3	1.51	10.43	20.26	41.01	72.96	30.58
$\sigma_K = h^*$	65.6	79.5	95.7	96.0	96.6	16.5	0.99	1.44	12.44	22.32	34.59	20.87
$\sigma_K = h_{ICI}^*$	66.1	94.4	95.8	96.0	96.5	1.6	1.14	1.34	1.55	11.67	27.57	10.33

Table 2: Results on the same data set with the local Gaussian approach proposed in [5]. Note that the last two lines make use of the scale section methods detailed in sec. 3.

For the sake of completeness and comparison, we present the performance of the approach presented in [5], which also use local image statistics modelled, however, by local Gaussian pdfs. Figure 4 and Table 2 show results for the local Gaussian case respectively equivalent to the one in Fig. 2 and Table 1 for the local Rayleigh. Although the DSC measure is not very discriminative between the two models, the MAD shows clearly that the Gaussian model is less competitive in comparison to the local Rayleigh one. This observation is also demonstrated by Fig. 4. Here too, the used of an automatic scale selection strategy improves significantly the segmentation results over the single scale segmentation algorithm. Finally, it is important to mention that in the Gaussian case there was a need to increase the contribution of the regularisation of the length term (in comparison to the Rayleigh case) in order to obtain acceptable results.

5 Conclusion

This paper presents the adaptation of the global model proposed by Sarti et al. [28], in order to use local image statistics. We kept the variational framework and the Rayleigh pdf to model the observe intensities of the ultrasound envelope image. We have also proposed an interesting adaptive scale selection procedure based on the ICI rule. Results on realistic simulations of ultrasound images show the robustness and the superiority of the proposed segmentation approach in comparison to [5, 28]. The efficiency and the genericity of the proposed scale selection strategy is also demonstrated.

10 BOUKERROUI: A LOCAL RAYLEIGH MODEL WITH SPATIAL SCALE SELECTIONS

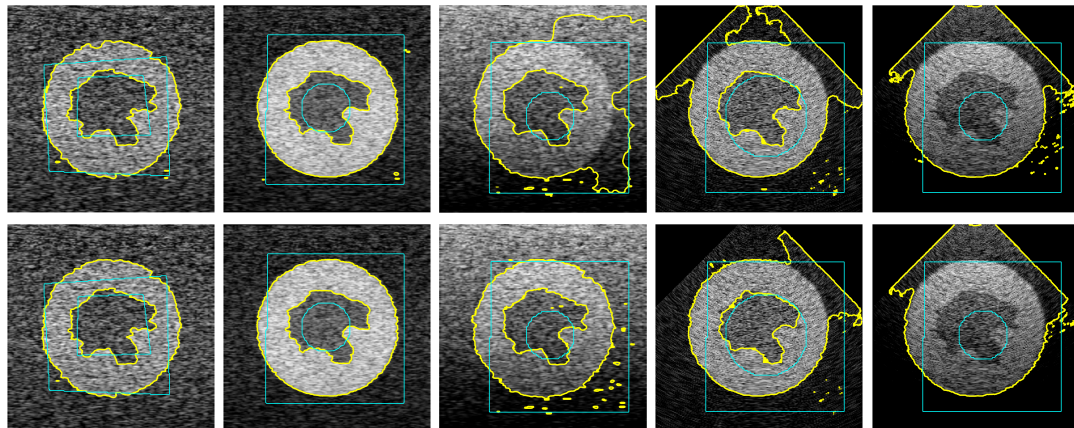


Figure 4: Same results as in Fig.2 but for the local Gaussian model. First line: scale selection of [27] with $\varepsilon = 0.5$. Second line: with the proposed ICI based approach with $\alpha = 0.01$.

References

- [1] E. A. Ashton and K. J. Parker. Multiple resolution bayesian segmentation of ultrasound images. *Ultrasonic Imag.*, 17(4):291–304, October 1995.
- [2] A. Belaid, D. Boukerroui, Y. Maingourd, and J.-F. Lerallut. Phase-based level set segmentation of ultrasound images. *IEEE Trans. Inf. Tech. Biomed.*, 15(1):138–147, 2011.
- [3] D. Boukerroui, O. Basset, A. Baskurt, and J.A. Noble. Segmentation of echocardiographic data. Multiresolution 2D and 3D algorithm based on gray level statistics. In *MICCAI*, pages 516–524, Cambridge, England, 1999. Springer-Verlag.
- [4] D. Boukerroui, A. Baskurt, J.A. Noble, and O. Basset. Segmentation of ultrasound images—multiresolution 2D and 3D algorithm based on global and local statistics. *Pattern Recognit. Lett.*, 24:779–790, 2003.
- [5] T. Brox and D. Cremers. On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional. *Int. J. Comput. Vis.*, 84(2):184–193, 2009.
- [6] C. B. Burckhardt. Speckle in ultrasound B-mode scans. *IEEE Trans. Sonics and Ultra.*, SU-25(1):1–6, 1978.
- [7] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim. A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Trans. Med. Imag.*, 15(3):290–298, June 1996.
- [8] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10(2):266–277, 2001.
- [9] V. Dutt and J. F. Greenleaf. Ultrasound echo envelope analysis using a homodyned K-distribution signal model. *Ultrasonic Imag.*, 16(4):265–287, October 1994.

- [10] T. Eltoft. The rician inverse gaussian distribution: A new model for non-rayleigh signal amplitude statistics. *IEEE Trans. Image Process.*, 14(11):1722–1735, November 2005.
- [11] T. Eltoft. Modeling the amplitude statistics of ultrasonic images. *IEEE Trans. Med. Imag.*, 25(2):1722–1735, February 2006.
- [12] J. Geusebroek, Arnold Smeulders, and Joost van de Weijer. Fast anisotropic gauss filtering. *IEEE Trans. Image Process.*, 12(8):938–943, 2003.
- [13] A. Goldenshluger and A. Nemirovski. On spatial adaptive estimation of nonparametric regression. *Math. Meth. Statistics*, 6:135–170, 1997.
- [14] M. F. Insana, R. F. Wagner, B. S. Garra, D. G. Brown, and T. H. Shawker. Analysis of ultrasound image texture via generalized Rician statistics. *Optical Engineering*, 25(6):743–748, 1986.
- [15] E. Jakeman. K-distributed noise. *J. Opt. A: Pure Appl. Opt.*, 1:784–789, 1999.
- [16] J. A. Jensen. Field: A program for simulating ultrasound systems. In *10th Nordic-Baltic Conf. on Biomedical Imag.*, volume 34, pages 351–353, 1996.
- [17] J. A. Jensen and N. B. Svendsen. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans. Ultrason., Ferroelec., Freq. Contr.*, 39:262–267, 1992.
- [18] Vladimir Katkovnik, Karen Egiazarian, and Jaakko Astola. Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule. *J. Math. Imag. Vis.*, 16(3):223–235, 2002.
- [19] S. Lankton and A. Tannenbaum. Localizing region-based active contours. *IEEE Trans. Image Process.*, 17(11):2029–2039, 2008.
- [20] F. Lecellier, J. Fadili, S. Jehan-Besson, G. Aubert, M. Revenu, and E. Saloux. Region-based active contours with exponential family observations. *J. Math. Imag. Vis.*, 36(1):28–45, 2010.
- [21] C. Li, C-Y Kao, J. C. Gore, and Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans. Image Process.*, 17(10):1940–1949, 2008.
- [22] N. Lin, W. Yu, and J. S. Duncan. Combinative multi-scale level set framework for echocardiographic image segmentation. *Med. Image Anal.*, 7:529–537, 2002.
- [23] M. Mulet-Parada and J. A. Noble. 2D+T acoustic boundary detection in echocardiography. *Med. Image Anal.*, 4(1):21 – 30, 2000. ISSN 1361-8415.
- [24] S. Nadarajah. Statistical distributions of potential interest in ultrasound speckle analysis. *Phys. Med. Biol.*, 52:213–227, 2007.
- [25] J. A. Noble and D. Boukerroui. Ultrasound image segmentation: A survey. *IEEE Trans. Med. Imag.*, 25(8):987–1010, 2006.
- [26] A. Papoulis and S. Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 4th edition, 2002.

12 *BOUKERROUI: A LOCAL RAYLEIGH MODEL WITH SPATIAL SCALE SELECTIONS*

- [27] J. Piovano and T. Papadopoulos. Local statistic based region segmentation with automatic scale selection. In *ECCV*, pages 486–499. Springer, 2008.
- [28] A. Sarti, E. Mazzini, C. Corsi, and C. Lamberti. Maximum likelihood segmentation of ultrasound images with rayleigh distribution. *IEEE Trans. Ultra. Fer. Freq. Control*, 52(6):974–960, June 2005.
- [29] P. M. Shankar, J. M. Reid, H. Ortega, C. W. Piccoli, and B. B. Goldberg. Use of non-Rayleigh statistics for the identification of tumors in ultrasonic B-scans of the breast. *IEEE Trans. Med. Imag.*, 12(4):687–92, 1993.
- [30] P. M. Shankar, V. A. Dumane, T. George, C W Piccoli, J M Reid, F Forsberg, and B B Goldberg. Classification of breast masses in ultrasonic B scans using Nakagami and K-distributions. *Phys. Med. Biol.*, 48(14):2229–2240, July 2003.
- [31] Z. Tao and H. Tagare. Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images. *IEEE Trans. Med. Imag.*, 25(11):1483–1491, Nov. 2006.
- [32] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez. Statistics of speckle in ultrasound B-scans. *IEEE Trans. Sonics and Ultra.*, 30(3):156–163, 1983.
- [33] L. Wang, L. He, A. Mishra, and C. Li. Active contours driven by local gaussian distribution fitting energy. *Singal Process.*, pages 2435–2447, 2009.
- [34] P. N. T. Wells. Ultrasound imaging. *Phys. Med. Biol.*, 51:R83–R98, 2006.
- [35] Q. Yang and D. Boukerroui. Optimal spatial adaptation for local region-based active contours: An intersection of confidence intervals approach. In *IMAGAPP*, pages 87–93, Algarve, Portugal, March 5-7 2011.
- [36] Y. Zhu, X. Papademetris, A.J. Sinusas, and J. S. Duncan. A coupled deformable model for tracking myocardial borders from real-time echocardiography using an incompressibility constraint. *Med. Image Anal.*, 14(3):429–448, 2010.

Efficient numerical schemes for gradient vector flow



Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Efficient numerical schemes for gradient vector flow

Djamal Boukerroui

Université de Technologie de Compiègne, CNRS, Heudiasyc UMR 6599, Centre de Recherche de Royallieu, BP 20529 – 60205 Compiègne Cedex, France

ARTICLE INFO

Article history:

Received 30 July 2010

Received in revised form

14 February 2011

Accepted 5 July 2011

Available online 18 July 2011

Keywords:

Active contours

Gradient vector flow

Numerical Schemes

ABSTRACT

Since its publication more than 10 years ago, the gradient vector flow (GVF) technique has been used and adapted to various models and problems. Its effectiveness has greatly contributed to its popularity. The main drawback of GVF and its generalisation, however, is their expensive computation load and its consequence on the capture range. In this work, we propose and compare different efficient numerical schemes to solve the GVF and its generalisations.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Active deformable contours are based generally on the minimisation of a two term energy function (internal and external) [1]. The internal energy encodes our prior about the shape of the contour and it preserves and imposes its smoothness. The external energy depends on image information and is designed to attract the solution to the desired boundary. Traditionally, the external energy is based only on edge information and is defined as $-f(\cdot)$, where $f(\cdot)$ is an edge map of the image $I(\cdot)$. The edge map is usually defined as the gradient norm of a Gaussian filtered version of the original image I , i.e. $|\nabla G_\sigma(\cdot) * I(\cdot)|^2$. Gradient vectors of the edge map image ∇f are normal to the edge, and pointing towards it. ∇f is known as the external potential force field that pulls the contour towards the desired edges. These gradient vectors have a large magnitude in the immediate vicinity of the boundary only and are still nearly zero in image locations away from the boundary. Such model encounters several difficulties such as, its sensitivity to initialisation due to noise and its low capture range, and more importantly its difficulties to converge to boundary concavities. Thus traditionally, close initialisation to the desired contour is often required.

Several researchers pointed out these limitations and have proposed alternative solutions among which, the balloon force [2,3] or the multi-resolution method [4]. These techniques extend the external force to a much larger range over the image domain and consequently reduce the snake's sensitivity to initialisation. Recent methods have also the ability to capture concavities and include the Charged-Particle Model (CPM) [5], the Charged Active

Contour based on Electrostatics (CACE) [6] and the Vector Field Convolution (VFC) [7]. In the CPM, the segmentation is performed by the simulation of charged particles moving in an electrostatic field. The classical potential force in active contours is replaced by electrostatic forces and the contour is modelled as a set of free moving particles. This modelling allows to model attractive and repulsive forces based on the charge sign of the particles. In [6], a new framework is proposed, introducing particle based electrostatics into active contours. The propagation is driven, as in the CPM, by a bi-directional force field based on the electrostatic interaction between an image-derived point charge field and a charged contour. The main difference between CACE and CPM is in the contour modelling, CACE is much faster and more efficient in convergence than CPM [6]. In [7], the external force field is calculated using a simple convolution of the edge map with a user-defined vector field kernel. This simplicity combined with its robustness to noise makes the VFC a very promising technique.

The most popular and first proposed solution is, however, the gradient vector flow [8] and its generalisation, the GGVF [9,10]. Since its publication about a decade ago, the GVF has been used and adapted to various models and problems (e.g. segmentation [8,11–14], tracking [15], segmentation with fuzzy spatial relations [13], registration [16], skeletonisation [17], interest point detection [18,19]). The popularity of Xu's GVF is certainly due to its effectiveness. A number of variants and extensions have been proposed in the literature. For instance, the Gradient and Direction Vector Flow (G&DVF) enhances the abilities of the GVF to converge to the desired boundary of complex shapes and prevents weak-edge leakage by adding prior directional information [14]. The Dynamic Directional GVF (DDGVF) [20] is an interesting variant, improved from the directional GVF approach [21]. The DDGVF snake is designed to be attracted to positive or negative

E-mail address: djamal.boukerroui@hds.utc.fr

step edges exclusively. The algorithm is particularly useful for the segmentation of object with complex boundaries [20]. Recently, the Hessian based Image Structure Adaptive GVF (HISAGVF) is proposed [22]. In this work, local image structure are taken into account, using the Hessian Matrix, in the diffusion term of the GVF. Therefore, the diffusion is anisotropic (constant for the GVF and nonlinear for the GGVF). The obtained force field improves convergence to very long concavity and preserves weak edge. The Normal-GVF restricts the diffusion to the normal direction of the edge [23]. This later work, guided Guillot and Guyader [24] to propose an original variant by reducing the number of unknowns to a single one, assuming that the expected vector field is the gradient field of a scalar function, and the diffusion is on the normal direction of the edge. Automatic initialisation methods of the GVF snake can be found in [25–27].

A major drawback of GVF and its generalisation GGVF is their expensive computation load and its consequence on the capture range. Indeed, as the GVF is an iterative diffusion method, the capture range of the resulting flow field depends highly on the number of iterations. Recently, the multigrid method was applied with success to speed up the computation [28,29]. In this paper, we propose and discuss other efficient alternatives.

The paper is organised as follows. A brief description of the mathematical formulation of the GVF is first given in Section 2. We review the explicit discretisation scheme proposed by Xu and Prince [8,10] and then present different efficient alternatives to solve the GVF problem and its generalisations. In Section 4, illustrative results are presented and discussed in order to demonstrate the new development. Section 5 summarises contributions. A preliminary version of this work was published in [30,31].

2. Background materials

The GVF approach extends the gradient map, ∇f , using a diffusion process. The GVF is defined as a vector field $\mathbf{v}(\mathbf{x}) : \Omega \rightarrow \mathbb{R}^m$ that minimises the following energy function [9]:

$$\mathcal{E} = \int_{\Omega} g(\mathbf{x}) |\nabla \mathbf{v}|^2 + h(\mathbf{x}) |\mathbf{v} - \nabla f|^2 d\mathbf{x}, \quad (1)$$

where $|\cdot|$ is the vector norm for tensor given by $\sqrt{\nabla \mathbf{v} \cdot \nabla \mathbf{v}}$. $g(\mathbf{x})$ and $h(\mathbf{x})$ are nonnegative functions on Ω . The basic idea behind this variational formulation is to make the solution smooth where there is no data, and to be as faithful as possible to the data where reliable measurements are possible. Using calculus of variations the solution of the above minimisation problem can be obtained using the following diffusion equation [9]:

$$\partial_t \mathbf{v} = \nabla \cdot (g(\mathbf{x}) \nabla \mathbf{v}) - h(\mathbf{x}) (\mathbf{v} - \nabla f), \quad (2)$$

with $\mathbf{v}(\mathbf{x}, t=0) = \nabla f$ as an initial state and reflective boundary conditions. Depending on the choices of the two weighting functions $g(\mathbf{x})$ and $h(\mathbf{x})$, Xu and Prince [8,9] proposed two vector flow fields.

(a) *Gradient vector flow*: The GVF is defined for the following special case:

$$g(\mathbf{x}) = \mu, \quad h(\mathbf{x}) = |\nabla f|^2, \quad (3)$$

where μ is a positive constant parameter governing the amount of the smoothness term in (1). Eq. (2) is then simplified to:

$$\partial_t \mathbf{v} = \mu \nabla^2 \mathbf{v} - h(|\nabla f|) (\mathbf{v} - \nabla f). \quad (4)$$

where ∇^2 is the Laplacian operator.

(b) *Generalised gradient vector flow*: In order to avoid smoothing in the proximity of detected edges and increase the amount of

regularisation in areas where information is constant, the following choices are adopted by Xu and Prince [9]:

$$g(\mathbf{x}) = e^{-|\nabla f|^2/\kappa}, \quad h(\mathbf{x}) = 1 - g(|\nabla f|), \quad (5)$$

where κ is a positive constant parameter. They decided, however, to approximate the evolution equation (2) by ignoring the variations on g . This defines the GGVF as the solution of

$$\partial_t \mathbf{v} = g(|\nabla f|) \nabla^2 \mathbf{v} - h(|\nabla f|) (\mathbf{v} - \nabla f). \quad (6)$$

We will denote by true generalised gradient vector flow (TGGVF), the solution of (2) using the weighting functions defined by (5). In the subsequent sections, we limit ourselves to two dimensional domains and give details only for the component u of the vector field \mathbf{v} in the x direction. From (6), we get:

$$\partial_t u = g \nabla^2 u - h u + h^1, \quad \text{with } h^1 = h \partial_x f. \quad (7)$$

3. Numerical implementations

Solving the problem numerically means finding discrete functions defined on a discrete grid which are a ‘good approximation’ of the continuous components of \mathbf{v} . We will denote by $u_{i,j}^k$ the discrete version of $u(i\Delta x, j\Delta y, k\tau)$ where τ is the time step and Δx and Δy are the spatial grid sizes. It is natural and hence common in the image processing community to use a uniform spatial grid (i.e. $\Delta x = \Delta y = h$). Finite difference methods are generally used in this context in order to approximate the different derivatives which lead to a discrete set of equations. For instance, the forward difference is used for the time derivative and central differences are used for spatial derivatives. This gives us the below approximations of $\partial_t u$ and of the Laplacian¹:

$$\partial_t u \approx \frac{u^{k+1} - u^k}{\tau},$$

$$\nabla^2 u|_{i,j} \approx \frac{u_{i+1,j} + u_{i,j+1} + u_{i-1,j} + u_{i,j-1} - 4u_{i,j}}{h^2}.$$

The substitution of the finite difference approximations in the evolution equation (7) results on different discrete iterative schemes, explicit, semi-implicit and implicit.

(a) *Explicit scheme*: The explicit scheme of the GGVF is given by [8]:

$$u_{i,j}^{k+1} = r_{i,j} (u_{i+1,j}^k + u_{i,j+1}^k + u_{i-1,j}^k + u_{i,j-1}^k - 4u_{i,j}^k) + (1 - \tau h_{i,j}) u_{i,j}^k + \tau h_{i,j}^1, \quad (8)$$

where $r = \tau h^{-2} g$. The above numerical scheme is called explicit as the solution at iteration $k+1$ is explicitly given, knowing the solution at the previous iteration k . The study of the stability of the above discrete schemes results to the following sufficient condition (see Appendix A)²:

$$\tau \leq \frac{2h^2}{\max_{i,j} (h^2 h_{i,j} + 8g_{i,j})}. \quad (9)$$

Explicit schemes are very popular for solving partial differential equations, mainly because of their simplicity. On the other hand,

¹ Although this approximation of the Laplacian is widely used by the computer vision community, this discretisation does not take into account its 2D nature as the approximation is not rotationally invariant. See [32] for an example of an isotropic discretisation.

² Our stability condition is tighter than Xu and Prince’s $\tau \leq \Delta x \Delta y / 4g_{\max}$ [9,8]. We believe that there was an error in Xu and Prince derivation. It is regrettable to notice that all works published about the GVF have reported the wrong stability condition. Fortunately the two conditions are equivalent for the GGVF and for the special case of the weighting functions given in (5) when $h = 1$.

this simplicity imposes very small and restrictive time steps to secure stability, leading to slow convergence. Such schemes are limited rather by their stability than by their accuracy [33]. Hence semi-implicit and implicit schemes are very desirable, as they remain numerically stable for larger time steps and therefore converge in less iterations. Such schemes can provide very interesting alternatives when the computational load is reduced using appropriate numerical methods (see e.g. [32,33]).

(b) *Semi-implicit scheme*: The implicit discretisation of (7) is given by:

$$\frac{u_{ij}^{k+1} - u_{ij}^k}{\tau} = \frac{g_{ij}}{h^2} (u_{i-1,j}^{k+1} + u_{i+1,j}^{k+1} + u_{i,j-1}^{k+1} + u_{i,j+1}^{k+1} - 4u_{ij}^{k+1}) - h_{ij} u_{ij}^{k+1} + h_{ij}^1. \quad (10)$$

Although this scheme is unconditionally stable (see Appendix A), the solution at each iteration is, however, obtained by solving a large linear system with a very sparse system matrix. The difficulty and the cost of resolution increases with the dimensionality of the domain Ω . We define the following semi-implicit scheme as:

$$\frac{u_{ij}^{k+1} - u_{ij}^k}{\tau} = \frac{g_{ij}}{h^2} (u_{i-1,j}^{k+1} + u_{i+1,j}^{k+1} + u_{i,j-1}^{k+1} + u_{i,j+1}^{k+1} - 4u_{ij}^{k+1}) - h_{ij} u_{ij}^k + h_{ij}^1, \quad (11)$$

where only one term in the RHS is left at time $k\tau$. The stability condition requires (see Appendix A):

$$\tau \leq \frac{2}{\max_{ij} h_{ij}}.$$

Here too we need to solve a linear system in order to obtain the solution at iteration $k+1$, in contrast to Eq. (10). Fortunately, the system matrix is data independent for the special case of GVF (meaning when $g_{ij} = \mu$). The Discrete Cosine Transform can be used to solve efficiently the semi-implicit scheme using the Fast DCT (the DCT respects the mirror boundary conditions). Considering a unit spatial grid size (i.e. $h=1$) and taking the two dimensional DCT of both sides of Eq. (11) leads us to:

$$\hat{u}_{m,n}^{k+1} = \frac{\hat{d}_{m,n}^k}{1 + 4\mu\tau \left(\sin^2\left(\frac{\pi m}{M}\right) + \sin^2\left(\frac{\pi n}{N}\right) \right)}, \quad (12)$$

where $\hat{d}_{m,n}^k = \text{DCT}\{(1 - \tau h_{ij})u_{ij}^k + \tau h_{ij}^1\}$ and $(m,n) \in [0, M-1] \times [0, N-1]$. Hence the calculation of the GVF field using the semi-implicit scheme requires 1 forward and 1 inverse DCT for each component per iteration.

(c) *Alternating direction explicit scheme*: The ADE is a two step method and is known to be unconditionally stable (see e.g. [34,35]). The scheme is obtained by using Saul'yev's approximations for the second order spatial derivatives used in the discretisation of the Laplacian term [36] as follows:

$$\nabla^2 u|_{ij} \approx \frac{u_{i-1,j}^{k+1} + u_{i,j-1}^{k+1} - 2u_{ij}^{k+1} - 2u_{ij}^k + u_{i+1,j}^k + u_{i,j+1}^k}{h^2},$$

$$\nabla^2 u|_{ij} \approx \frac{u_{i-1,j}^k + u_{i,j-1}^k - 2u_{ij}^k - 2u_{ij}^{k+1} + u_{i+1,j}^{k+1} + u_{i,j+1}^{k+1}}{h^2}.$$

Thus, replacing the Laplacian term in the implicit scheme given in Eq. (10) by the above discrete approximations, leads to the following two steps discretisation of Eq. (7):

$$u_{ij}^{k+1} - \frac{r_{ij}(u_{i-1,j}^{k+1} + u_{i,j-1}^{k+1})}{1 + 2r_{ij} + \tau h_{ij}} = \frac{(1 - 2r_{ij})u_{ij}^k}{1 + 2r_{ij} + \tau h_{ij}} + \frac{r_{ij}(u_{i+1,j}^k + u_{i,j+1}^k)}{1 + 2r_{ij} + \tau h_{ij}} + \frac{\tau h_{ij}^1}{1 + 2r_{ij} + \tau h_{ij}}, \quad (13)$$

$$u_{ij}^{k+2} - \frac{r_{ij}(u_{i+1,j}^{k+2} + u_{i,j+1}^{k+2})}{1 + 2r_{ij} + \tau h_{ij}} = \frac{(1 - 2r_{ij})u_{ij}^{k+1}}{1 + 2r_{ij} + \tau h_{ij}} + \frac{r_{ij}(u_{i-1,j}^{k+1} + u_{i,j-1}^{k+1})}{1 + 2r_{ij} + \tau h_{ij}} + \frac{\tau h_{ij}^1}{1 + 2r_{ij} + \tau h_{ij}}. \quad (14)$$

Although it is not obvious, the above scheme is fully explicit. Indeed, in Eq. (13), the calculation proceeds in a raster scan and Eq. (14) proceeds in an inverse raster scan. This numerical method suffers from anisotropy when very large time steps are used. In order to reduce such effect, left–right flip of the domain is recommended between iterations. We will denote by ADES this variant.

(d) *Operating splitting schemes*: Inspecting the evolution equation (2), we notice that the diffusion term takes exactly the form of the one in nonlinear diffusion filtering ([33], Eq. (1)). Splitting methods were used successfully to solve the later problem. In our case, there is an additional extra term “ $h(\mathbf{x})(\mathbf{v} - \nabla f)$ ”. Such techniques are known as Approximation Factorisation methods [37]. For notational convenience, we rewrite the implicit scheme of (2) in a matrix-vector notation:

$$(\mathbb{I} - \tau A_x - \tau A_y)u^{k+1} = u^k + \tau h^1, \quad (15)$$

where \mathbb{I} is the unit matrix and $u \in \mathbb{R}^m$, $m=2$, is represented as a vector. A matrix A_l corresponds to the derivatives along the l^{th} coordinate axis, i.e. the matrix-vector multiplication $A_l u$ is the discrete approximation of the following linear operator:

$$L_{x_l}(u) = \partial_{x_l}(g(\mathbf{x})\partial_{x_l}u) - \frac{1}{m}h(\mathbf{x})u, \quad (16)$$

simply given by [33,37]

$$\sum_{j \in \mathcal{N}(i)} \frac{g_j + g_i}{2h^2} (u_j - u_i) - \frac{1}{m}h_i u_i, \quad (17)$$

where $\mathcal{N}(i)$ is the set of two neighbours of pixel i along the l^{th} axis (boundary pixel have only one neighbour). Thus, the elements of A_l are given by

$$(a_l)_{ij} = \begin{cases} \frac{g_j + g_i}{2h^2} & (j \in \mathcal{N}(i)), \\ -\frac{h_i}{m} - \sum_{n \in \mathcal{N}(i)} \frac{g_i + g_n}{2h^2} & (j = i), \\ 0 & (\text{else}). \end{cases} \quad (18)$$

Barash et al. [37] analysed different splitting operator to solve the diffusion equation that could be used to solve Eq. (15). Namely the Additive Operator Splitting (AOS):

$$u^{k+1} = \frac{1}{2} \sum_{l \in \{x,y\}} (\mathbb{I} - 2\tau A_l)^{-1} (u^k + \tau h^1), \quad (19)$$

the Locally One Dimensional (LOD):

$$u^{k+1} = \prod_{l \in \{x,y\}} (\mathbb{I} - \tau A_l)^{-1} (u^k + \tau h^1) \quad (20)$$

and the Additive Multiplicative Operator Splitting (AMOS):

$$u^{k+1} = \frac{1}{2} [(\mathbb{I} - \tau A_x)^{-1} (\mathbb{I} - \tau A_y)^{-1} + (\mathbb{I} - \tau A_y)^{-1} (\mathbb{I} - \tau A_x)^{-1}] (u^k + \tau h^1). \quad (21)$$

The LOD is known to be the most efficient, but not symmetric, method. The AOS has the advantage of being symmetric and the AMOS is the symmetrised version of the LOD. All the above approximations, however, do not account for the Factorisation Error, i.e. the difference between the true operator in (15) and the approximation (see e.g. [38]). Tested on our problem, all produced reasonably good results for small time steps. Analysing the factorisation error of the LOD approximation, we obtain:

$$E_{AF} = \tau^2 A_x A_y u^{k+1}. \quad (22)$$

It is a second time order error and involves mixed forth-order derivatives of the solution at step $k+1$. This explains one of our observations about the error behaviour of the LOD scheme. The LOD method has the tendency to misestimate the gradient vector field near edge locations. As an alternative, we propose the below approximation factorisation:

$$(\mathbb{I} - \tau \mathbb{A}_x)u^* = (\mathbb{I} + \tau \mathbb{A}_y)(u^k + \tau h^1)$$

$$(\mathbb{I} - \tau \mathbb{A}_y)u^{k+1} = u^* - \tau \mathbb{A}_y u^k \quad (23)$$

with a factorisation error of:

$$E_{AF} = \tau^2 \mathbb{A}_x \mathbb{A}_y (u^k - u^{k+1}). \quad (24)$$

In contrast to the LOD scheme, the error depends on the residual and not on the solution. Hence, we expect it to decrease as the method converges. We denoted this scheme by the AFI scheme.

Given the above form of \mathbb{A}_i , all the above approximation schemes can be solved efficiently by the Thomas algorithm, as proposed in [33]. Notice that the above iterative solutions (19)–(23) solve the TGGVF and without any extra computational cost, as the matrices \mathbb{A}_i are calculated only once.

4. Results and discussion

We have presented several alternatives to the basic explicit scheme that allow larger time steps. All schemes are first order in time and have a second order accuracy in space and are unconditionally stable (apart from the semi-implicit scheme). All methods have an $O(N^2)$ complexity, except for the semi-implicit scheme which uses a fast DCT based solver with a complexity of $O(N^2 \log_2 N)$, for an image of size $N \times N$. The computational time for one iteration of the ADE schemes is roughly the same as for the explicit scheme, and within sight of 2–3 times slower for the ADES, LOD, AOS and AFI schemes. The AMOS runs about five times slower. This results are based on an average of 600 iterations runs and for eight different image sizes. This imposes a minimum time step between 0.25 and 1.5 for the proposed methods to be faster than the explicit one. It is important to recall, however that the use of large time steps may introduce a loss of accuracy, and a loss of isotropy for some of the proposed schemes. In this section, we propose a way to quantify their accuracy when larger time steps are used.

Because of the restrictive stability condition of the semi-implicit scheme, generally a maximum time step between 3 and 5, the proposed DCT based solver is only about 2–3 times faster than the explicit scheme. Moreover, the DCT based solver cannot be applied to the GGVF. It is worthy to point out though that such method may be interesting if a higher order of accuracy in space is required as its computational burden is independent of the finite difference approximation of the Laplacian operator. This

method has been included from an exhaustive perspective and we will not discuss it further.

4.1. Comparison criteria

In order to quantify the accuracy of the different numerical schemes, we also have to keep in mind our final objective for the use of GVF. Indeed, in most usage of the GVF technique, the important information is the orientation of the vectors. For example, the amplitude is rarely used in segmentation applications. To this end, we synthesised an edge map test image composed of a perfect circle for which we could compute the theoretical optimal orientation map of the ideal vector field (see Fig. 1). Two criteria are then used to evaluate accuracy: the value of the energy function \mathcal{E} given by Eq. (1), and the mean absolute angular error of the estimated vector field in comparison to the ground truth. Following [39], we define the angular error between two vector fields (u_c, v_c) and (u_e, v_e) as the arithmetic mean of

$$\arccos\left(\frac{u_c u_e + v_c v_e + \varepsilon}{\sqrt{(u_c^2 + v_c^2 + \varepsilon)(u_e^2 + v_e^2 + \varepsilon)}}\right), \quad (25)$$

where ε is a small positive constant.

4.2. Comparison

The first experiment uses the ideal noise free test image. The results for three different values of the time step τ are shown in Fig. 2. Results of the explicit scheme, obtained with a time step of 0.2, are also shown for comparison.

The first observation is the expected behaviour of all methods for being very close to the explicit one for small τ values (not shown) and then divert as τ increases. An important observation is that the proposed measure is very informative as its behaviour is not always correlated with the energy (see for example the behaviours of the AFI and the AOS schemes). Notice for example the linear decrease, in a logarithmic scale, of the angular error of the DCT scheme. The decrease is correlated to the extension of the progress of the diffusion process, therefore to the capture range of the obtained vector field.

We notice that, for τ values less than 10, the curves of the angular errors are within 1° for the ADES and AFI schemes, and within 2° for all methods except for LOD and AOS. The worst behaviours are obtained by the LOD and the AOS methods (the highest error reaches 5° at time $t=300$ for the LOD method with $\tau=25$). This confirms the observations made by Papandreou and Maragos when the AOS scheme was used in active contour models [40].

The results of the proposed AFI and ADES are highly satisfactory. The angular errors are of the same order as the explicit method for $t > 50$ with $\tau=10$ and $t > 150$ for $\tau=25$; in other words only after 4–6 iterations. As expected, the AFI method

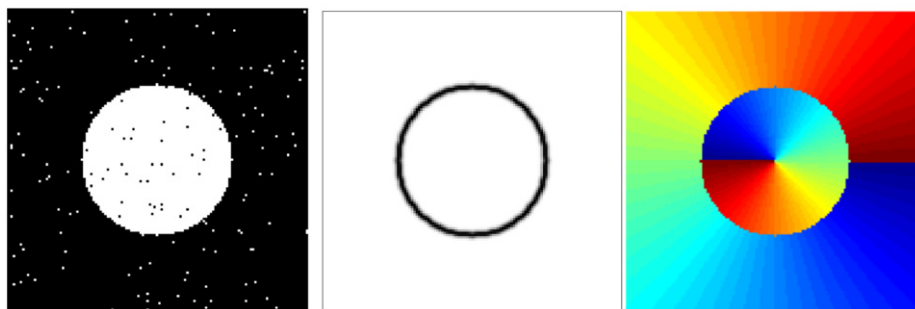


Fig. 1. Test image with 2% impulse noise (left). Example of a noise free edge map test image obtained with gaussian filtering $\sigma=2.5$ (middle). Theoretical optimal orientation map of the ideal vector field for the noise free image (right). Image size 129×129 .

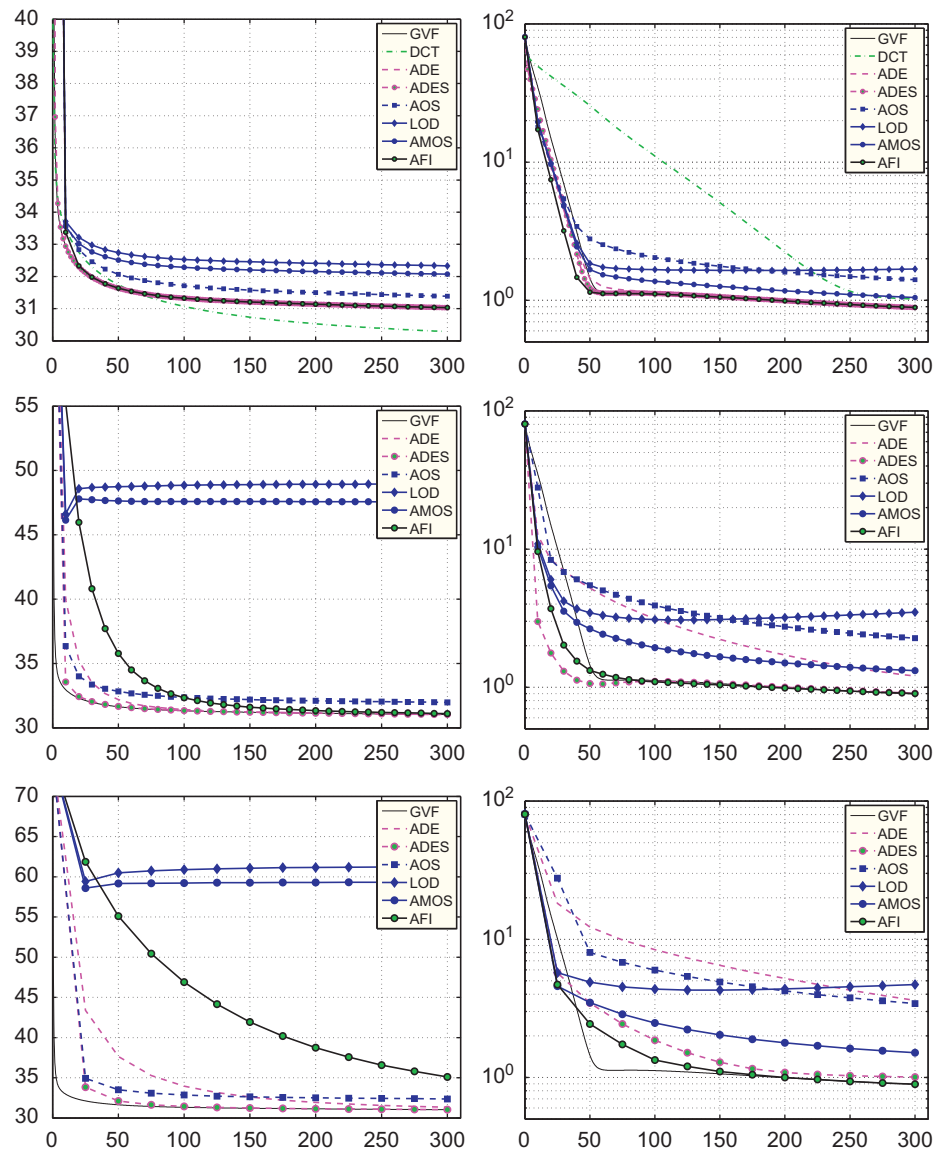


Fig. 2. (left) Evolution of the GGVF energy function (1) versus time for different schemes; (right) Evolution of the mean absolute angular error in degree versus time. First line: $\tau = 2$; Second line: $\tau = 10$; Third line: $\tau = 25$. Results on the noise free edge test image obtained with the GGVF for $\kappa = 0.2$. The DCT scheme was shown for comparison with the GVF energy.

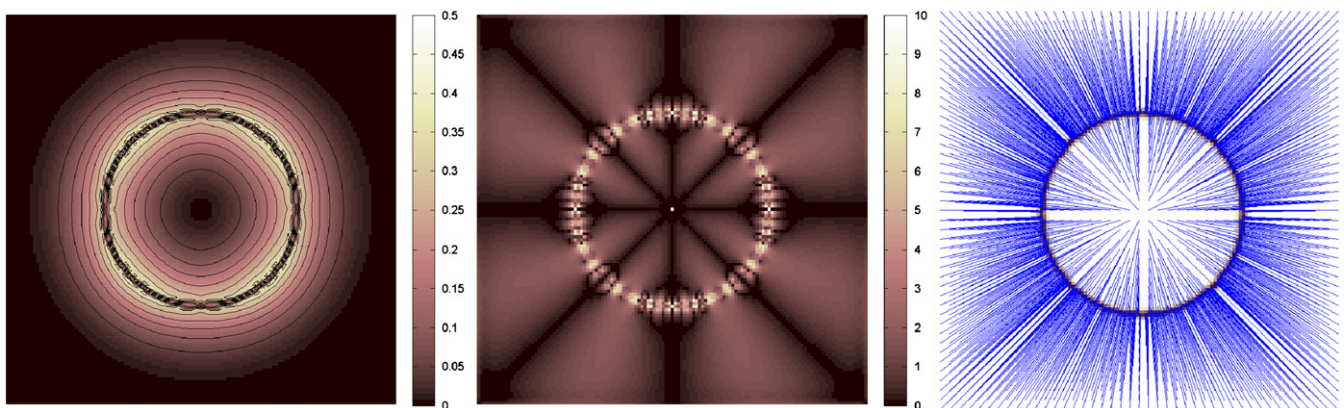


Fig. 3. From left to right: Image of the norm of the GVF field obtained after 300 iterations of the explicit scheme with $\tau = 0.2$, $\mu = 0.2$; isocurves were also shown. Image of the absolute angular error of the result of the explicit scheme (in degree). An example of streamlines generated from the obtained GVF field.

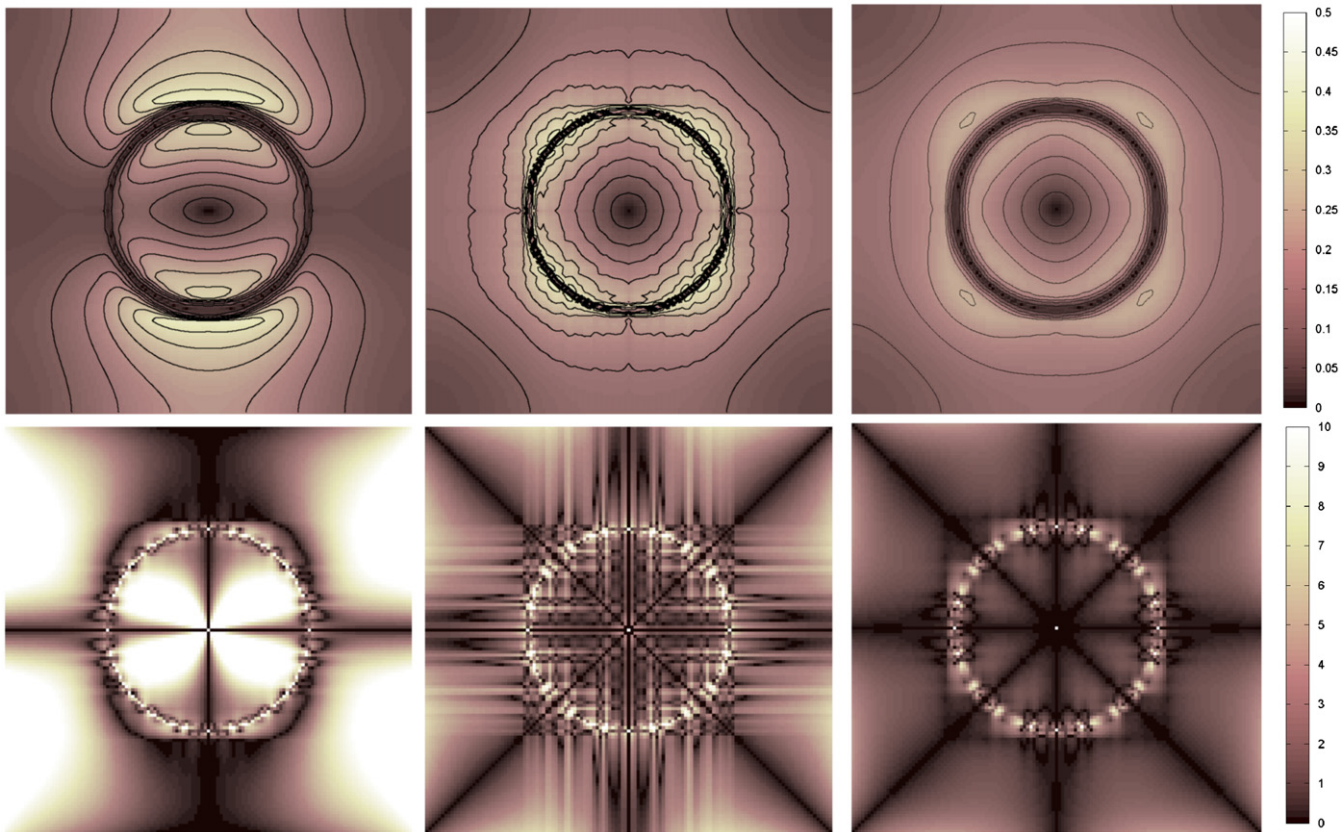


Fig. 4. First line: Image of the norm of the GVF with isocurves. Second line: Image of the absolute angular error (in degree). Respectively from the left to the right column: LOD, AOS and AMOS scheme. All results are obtained on the edge map of the noise free image with $\sigma = 1$, $\mu = 0.2$, $\tau = 25$, after 25 iterations.

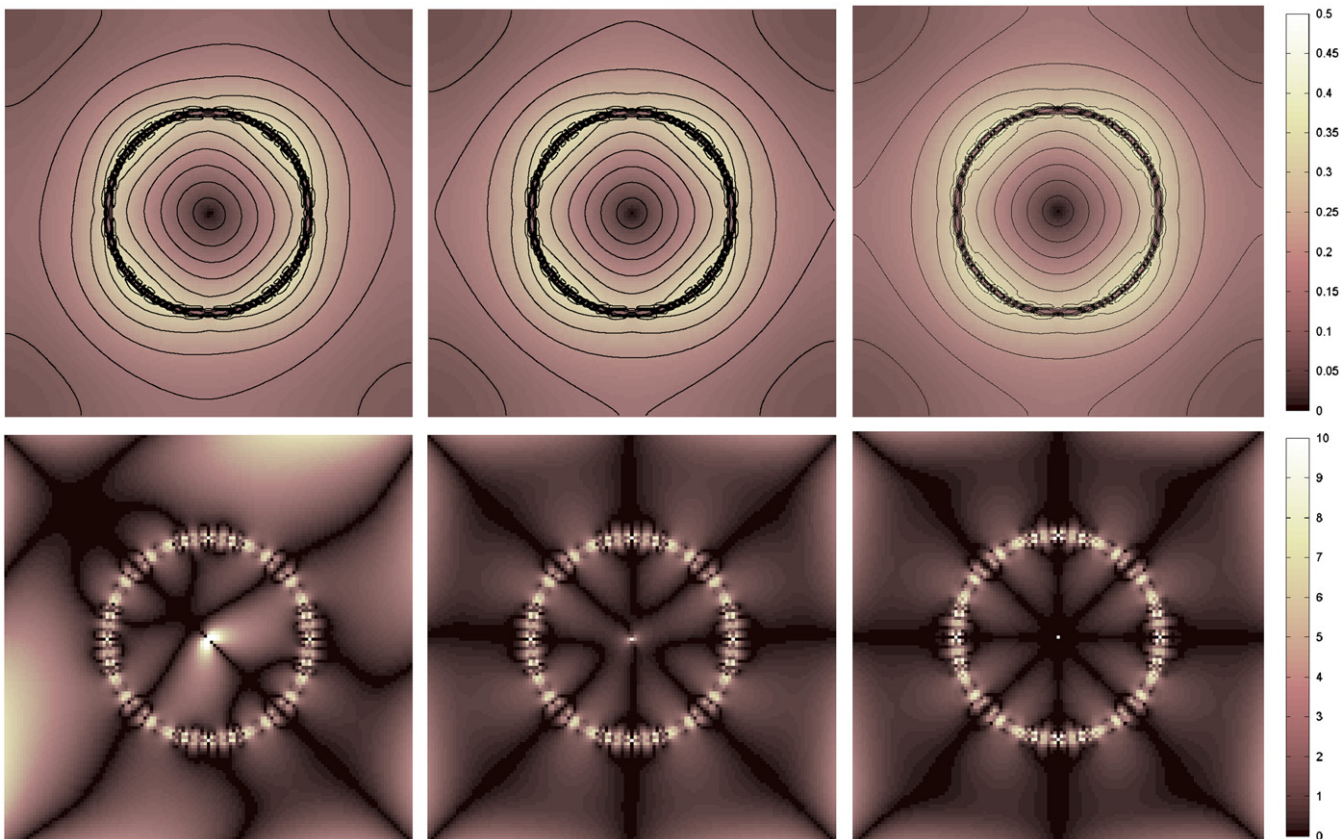


Fig. 5. Same results as in Fig. 4 for the ADE, ADES and AFI schemes respectively from the left to the right column.

improves considerably the performances of approximation methods. It showed, however, a lower rate of convergence in terms of minimising (1). The ADES method gives the best results in this experiment. Indeed, notice how close the ADES curves are to the GVF ones.

In support of the analysis, Figs. 4 and 5 show the amplitude of the obtained fields for the different proposed methods. Equivalent images for the explicit scheme are shown in Fig. 3 for comparison. These images give us a visual assessment of the extent of the capture range. The isocurves are also presented in order to show locally the direction of the vectors. Images of the angular error are

shown in the interest of assessing the isotropy of the obtained vector field. Notice the imprecision of the AOS scheme and the dissymmetry (loss of isotropy) of the LOD and ADE schemes. The results of the AMOS are acceptable in terms of the decrease of the angular error, but not in terms of energy minimisation and at the expense of a double computational load.

In order to quantify the loss of accuracy when using larger time steps and to study the influence of the diffusion parameters on accuracy, we compared the vector field \mathbf{v} , obtained with the different schemes, to the vector field \mathbf{v}_0 , estimated using the explicit one after a diffusion time $t = 200$. We used the previously

Table 1

Comparison of estimated vector fields \mathbf{v} with the different proposed methods to the one obtained with the explicit scheme \mathbf{v}_0 , with $\tau = 0.2$, taken as ground truth. The table shows the average angular errors, θ_E , as define in Eq. (25), and the relative l_2 error norms in percentage. The noise free image shown in Fig. 1 is used in this experiment; The edge map image is obtained with Gaussian filtering with $\sigma = 1$. Diffusion time $t = 200$, equivalent to 1000 iterations of the explicit scheme. The minimum and the maximum values are highlighted.

Noise free image												
Method	τ	Iter.	AOS		LOD		AMOS		AFI		ADES	
			θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$
GVF, $\mu = 0.1$	5	40	<u>3.9</u>	5	2.2	<u>13</u>	2.0	8	2.6	<u>0</u>	<u>0.3</u>	<u>0</u>
	15	14	<u>8.5</u>	8	5.8	<u>27</u>	5.5	17	6.2	<u>1</u>	<u>2.8</u>	2
	25	8	<u>10.6</u>	11	7.2	<u>34</u>	6.8	23	7.5	3	<u>4.4</u>	<u>1</u>
	35	6	<u>12.8</u>	11	9.4	<u>39</u>	8.8	27	9.3	5	<u>7.7</u>	<u>1</u>
GGVF, $\kappa = 0.1$	5	40	1.0	4	<u>2.6</u>	<u>29</u>	0.3	10	0.0	1	<u>0.0</u>	<u>0</u>
	15	14	2.2	6	<u>4.3</u>	<u>46</u>	0.6	17	0.1	2	<u>0.1</u>	<u>2</u>
	25	8	3.2	10	<u>4.9</u>	<u>51</u>	0.9	22	<u>0.4</u>	6	0.5	<u>4</u>
	35	6	3.9	11	<u>5.4</u>	<u>55</u>	1.1	24	<u>0.6</u>	8	1.0	<u>7</u>
GGVF, $\kappa = 0.2$	5	40	1.2	9	<u>1.9</u>	<u>23</u>	0.4	11	0.0	1	<u>0.0</u>	<u>0</u>
	15	14	2.4	11	<u>3.3</u>	<u>37</u>	0.8	18	0.2	2	<u>0.1</u>	<u>2</u>
	25	8	3.5	15	<u>3.9</u>	<u>42</u>	1.0	24	0.6	6	<u>0.5</u>	<u>4</u>
	35	6	4.2	15	<u>4.5</u>	<u>46</u>	1.2	26	<u>0.8</u>	9	1.0	<u>7</u>
GGVF, $\kappa = 0.5$	5	40	<u>1.4</u>	12.8	1.2	<u>15.0</u>	0.5	10.5	0.1	<u>0.5</u>	<u>0.0</u>	<u>0.5</u>
	15	14	<u>2.8</u>	16.6	2.3	<u>26.8</u>	0.9	17.9	0.5	5.0	<u>0.1</u>	<u>1.4</u>
	25	8	<u>3.9</u>	20.8	3.0	<u>32.1</u>	1.2	23.9	1.2	13.1	<u>0.5</u>	<u>5.4</u>
	35	6	<u>4.6</u>	21.8	3.5	<u>36.3</u>	1.4	26.3	1.6	22.9	<u>1.1</u>	<u>8.6</u>

Table 2

Same results as in Table 1 for the image shown in Fig. 1 with 2% impulse noise.

Noisy image												
Method	τ	Iter.	AOS		LOD		AMOS		AFI		ADES	
			θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$	θ_E°	$\mathcal{E}(\%)$
GVF, $\mu = 0.1$	5	40	<u>4.0</u>	7	2.6	<u>12</u>	2.3	8	1.9	1	<u>1.7</u>	<u>1</u>
	15	14	<u>8.5</u>	11	4.4	<u>25</u>	3.9	17	2.9	<u>2</u>	<u>2.5</u>	<u>2</u>
	25	8	<u>12.4</u>	15	5.2	<u>31</u>	4.3	23	5.0	5	<u>2.0</u>	<u>1</u>
	35	6	<u>15.6</u>	18	6.5	<u>36</u>	5.6	27	9.1	10	<u>2.8</u>	<u>2</u>
GGVF, $\kappa = 0.1$	5	40	<u>9.7</u>	11	5.7	<u>31</u>	3.9	17	<u>1.9</u>	<u>1</u>	<u>1.9</u>	<u>1</u>
	15	14	<u>17.5</u>	19	9.5	<u>48</u>	6.6	30	8.1	7	<u>2.1</u>	<u>1</u>
	25	8	<u>21.8</u>	24	11.5	<u>54</u>	8.2	36	13.9	17	<u>3.8</u>	<u>3</u>
	35	6	<u>24.4</u>	27	13.3	<u>58</u>	9.6	41	17.2	24	<u>7.8</u>	<u>6</u>
GGVF, $\kappa = 0.2$	5	40	<u>7.5</u>	12	4.0	<u>24</u>	2.7	14	<u>1.9</u>	<u>1</u>	<u>1.9</u>	<u>1</u>
	15	14	<u>14.7</u>	18	6.8	<u>38</u>	4.1	25	8.1	8	<u>2.4</u>	<u>2</u>
	25	8	<u>19.4</u>	23	8.0	<u>44</u>	<u>4.9</u>	31	14.1	19	6.0	<u>4</u>
	35	6	<u>22.4</u>	26	9.3	<u>48</u>	<u>5.9</u>	36	17.1	27	11.9	<u>9</u>
GGVF, $\kappa = 0.5$	5	40	<u>5.2</u>	14.4	2.8	<u>15.4</u>	2.5	11.7	<u>2.0</u>	1.7	<u>2.0</u>	<u>1.6</u>
	15	14	<u>10.9</u>	21.8	4.4	<u>27.2</u>	<u>3.1</u>	20.9	9.0	12.4	4.5	<u>2.5</u>
	25	8	<u>15.8</u>	27.1	5.3	<u>32.8</u>	<u>3.5</u>	26.8	15.0	30.0	11.7	<u>10.0</u>
	35	6	<u>19.0</u>	30.4	6.3	37.0	<u>3.9</u>	30.4	17.7	<u>43.6</u>	21.8	<u>23.1</u>

defined angular error, and the relative l_2 norm error, defined as:

$$\mathcal{E} = \frac{\sqrt{(\mathbf{v} - \mathbf{v}_0) \cdot (\mathbf{v} - \mathbf{v}_0)}}{\sqrt{\mathbf{v}_0 \cdot \mathbf{v}_0}} \quad (26)$$

The results on the noise free image are summarised in Table 1. This experiment shows clearly that the ADES and the AFI schemes outperform the other approximation factorisation methods. The accuracy of the ADES method is acceptable even for large time steps, for all parameter combinations. As predicted, AMOS is in general better than LOD and AOS. An interesting observation is the opposite behaviour of the LOD, in comparison to other schemes, when the diffusion term is reinforced by larger values of κ in the GGVF. We have also implemented the Full MultiGrid solver, proposed in [29] to solve the GGVF problem. The application of the FMG-GS(1,2,2) scheme produces an angular error of 1.9° for $\kappa = 0.2$. The ADES error is lower than 1.9° after four iterations with $\tau = 15$.

4.2.1. Influence of noise

Finite difference methods can be very sensitive to noise. To test the noise influence on the loss of accuracy when using larger time steps, we performed the same experiments presented in Table 1, with 2% of impulse noise. The test image is shown in Fig. 1, and the results are shown in Table 2. Notice that it is the original image that is corrupted by the impulse noise and not the edge map image. Thus, the interpretation of the results can be twofold: either the isolated pixels are considered as noise and therefore the vector field should be smoother or they considered as weak edges and the diffusion should be stopped in order to preserve the boundaries. To help interpretation of the results, the

flow fields of the explicit scheme, considered as the ground truth in this experiment, are shown in Fig. 6. It can be observed that the diffusion in the GGVF is proportional to κ .

Analysing Table 2, it can be observed that the ADES scheme is by far the closest method to the explicit scheme for low diffusion fields (GVF, GGVF, $\kappa = 0.1$). It is also the closest method for small τ values when the diffusion is increased (i.e. large κ values). The accuracy of the AFI scheme is more affected by the increase of τ and its behaviour approaches the AOS scheme for $\kappa = 0.5$. The AMOS gives acceptable results in terms of the angular error for large time steps for $\kappa = 0.2$ and 0.5 . The AOS scheme has the highest angular errors whereas the LOD has the highest amplitude errors. Some of the obtained flow fields for the AOS, AMOS, AFI and ADES are shown in Fig. 7. The selection is made to better illustrate the differences in comparison to the results of the explicit scheme. For example, the figure shows that the AOS scheme stops the diffusion even at isolated pixel boundaries. The AFI has the same drawback but only for larger time steps. The AMOS produces smooth vector field but has a larger amplitude errors. The strong diffusion may lead to boundary leakage in segmentation application and the amplitude errors may be problematic in other applications.

4.3. Discussion

The results of the Full MultiGrid solver on this data are shown in Fig. 6. We first notice that the FMG method produces generally smoother fields, than the explicit scheme. To some extent, this is of course to be expected as in the explicit scheme, the diffusion is stopped after a certain time. The cause also might be a parameter

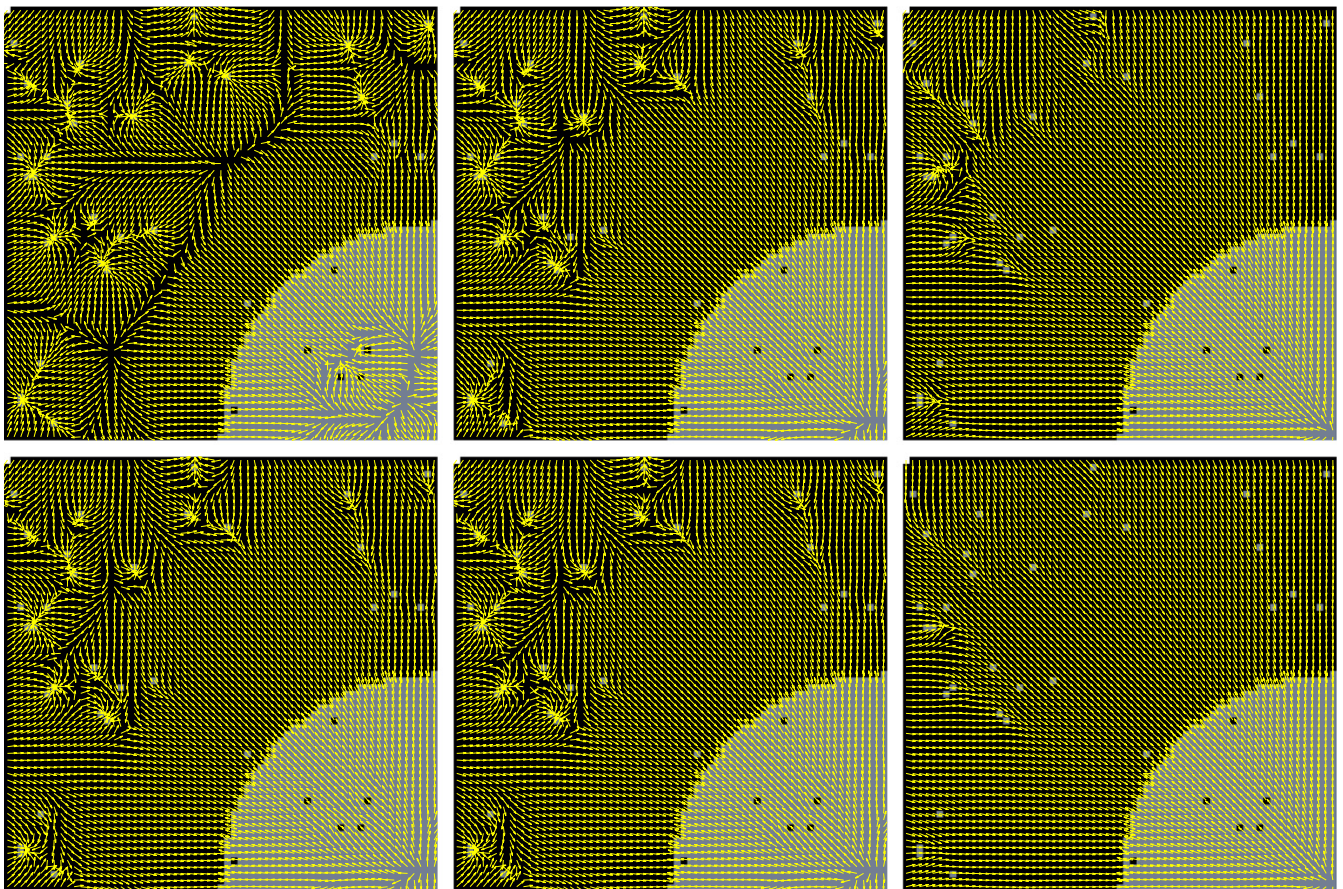


Fig. 6. Vector field obtained by the explicit scheme after 1000 iterations (first row), the FMG-GS (1,2,2) (second row). Original image shown in Fig. 1. From left to right: result of the GVF, $\mu = 0.1$; GGVF $\kappa = 0.1$ and 0.5 . Only the top right corner of the image is shown.

tuning effect of the FMG method. An example is the result presented in Fig. 6, for the GVF (first column). Here, the choice of the diffusion parameter μ was made in order to not smooth out all the pixels due to the impulse noise as they may be seen as

weak boundaries. The FMG solver produced an over smoothed solution. The middle column is an intermediate case where both method produced similar results. In the right column, the desired solution should not to be influenced by the impulse noise.

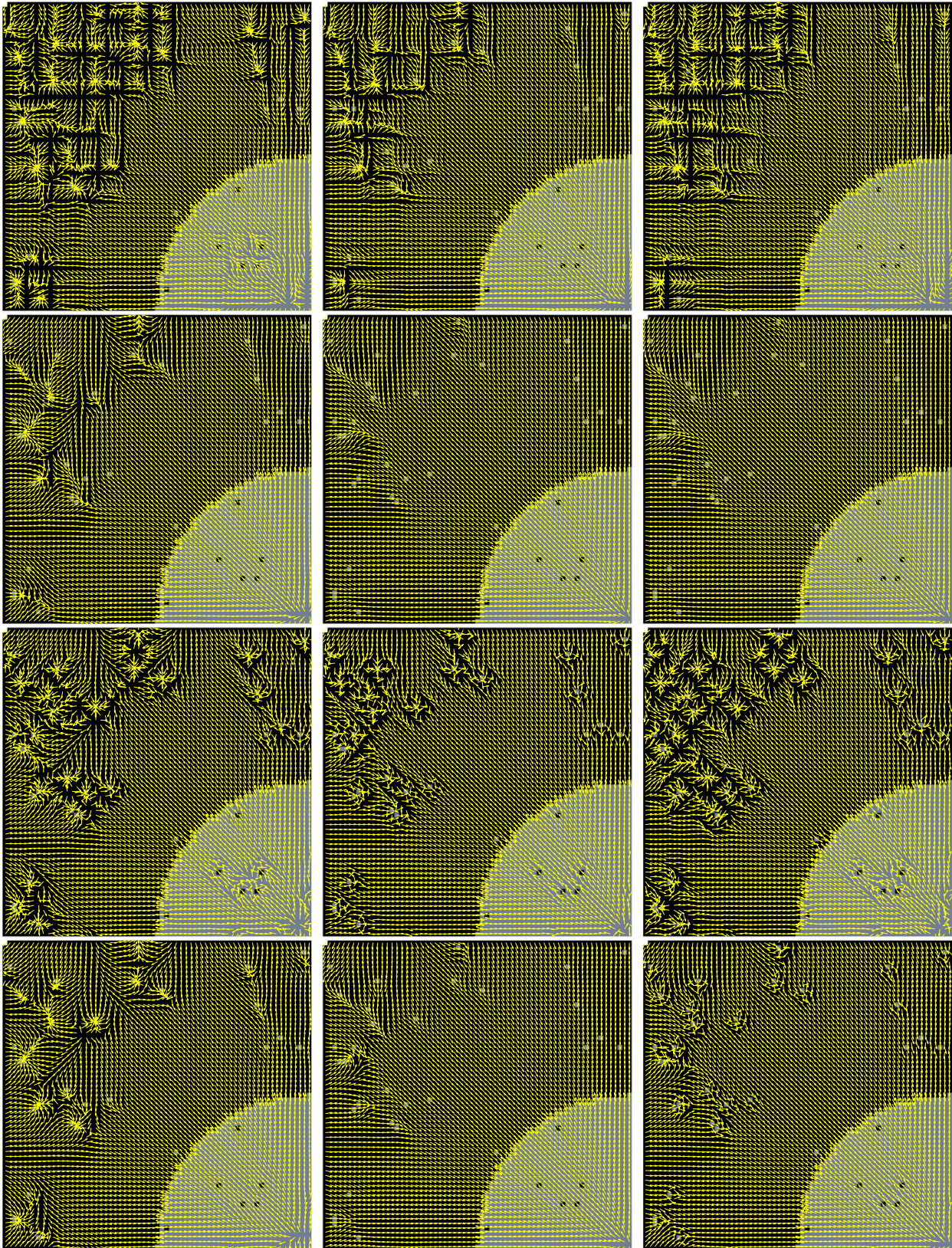


Fig. 7. From top to bottom row: GGVF fields obtained by the AOS, AMOS, AFI and ADES. $\kappa = 0.1$ and $\tau = 25$ (left column); $\kappa = 0.5$ and $\tau = 15$ (middle column); $\kappa = 0.5$ and $\tau = 35$ (right column). The corresponding fields obtained by the explicit scheme and the FMG method are shown in Fig. 6.

The FMG solver produced a slightly smoother field, therefore preferable than the explicit scheme. The corresponding results for the GGVF obtained by the other methods are shown in Fig. 7.

In iterative diffusion techniques, the stopping time has a strong affect on the diffusion results [41]. Although, Xu and Prince [42] have shown the convexity, and therefore the global optimality of the GVF energy function, it is unclear, how smooth or close to the data is the optimal solution. Therefore, the stopping time is an important tuning parameters, that may offer more flexibility than the diffusion parameters μ or κ . A first straightforward adaptation of [41], for the GVF problem appeared in [43].

Our main point here is the following. Multigrid techniques are known to be efficient PDE solvers. Therefore, the FMG solver is probably the best method, both in terms of accuracy and efficiency, to find a close solution to the global optimum of minimising (1). Is the optimal solution of (1), however, the desirable solution? Iterative techniques offer an extra parameter, the stopping time, to find the most desirable solution. In this perspective, the proposed ADES scheme is an interesting alternative, given the fact that in most cases, the ADES method has produced the closest results to the explicit scheme when larger time steps are used. A speed factor of about 60 can be achieved without important losses of accuracy and isotropy.

5. Summary and conclusion

Since its introduction by Xu and Prince, gradient vector flow has become very popular and is extensively used in contour based segmentation as an external energy force. Probably the main drawback of such a force is its high computation burden. Efficient schemes for the minimisation of the GVF and its generalisation are proposed and analysed. We have shown that the ADES scheme may be a good alternative to the multigrid technique. The proposed AFI scheme also shows a better behaviour than the LOD and AOS approximation factorisation methods. Although included, we did not discuss much the DCT based solver as it is only appropriate for the GVF and because computationally it is the most expensive.

Finally we may emphasis the fact that the Alternative Direction Explicit schemes are not well known by the image processing community. They may be of interest to solve other image processing PDE problems.

Acknowledgements

This work was partially funded by the regional research project ECHOPEDIA (région Picardie, France). The author thanks Izumi Otani for English proofreading.

Appendix A. Stability of a difference scheme

The stability of a difference scheme can be studied by means of the discrete von Neumann criterion for stability (see e.g. [32]). This approach considers a discrete Fourier mode for the problem. In our case given by:

$$u_{ij}^n = \xi^n \exp \{i\pi h(ik+jl)\},$$

where $i = \sqrt{-1}$ and the subscript on the ξ term is a multiplicative exponent. Inserting this general Fourier mode in a difference scheme, a necessary condition for stability is obtained by restricting h and τ so that $|\xi| \leq 1$. Note that it is easy to show that the stability of a difference system of the form $u^{n+1} = \rho u^n + c$ is

equivalent to the stability of $u^{n+1} = \rho u^n$ subject to the condition that c is bounded. Hence we will not be considering the constant term τh_{ij}^1 in Eqs. (8), (10) and (11).

- (a) *Explicit scheme*: Doing so for the scheme given in Eq. (8) we obtain:

$$\xi = [1 - \tau h_{ij} - 4\tau r_{ij} + 2\tau r_{ij}(\cos(\pi kh) + \cos(\pi lh))]. \quad (A.1)$$

Let us denote by $\gamma = \cos(\pi kh) + \cos(\pi lh)$ and replace τr_{ij} by $\tau h^{-2} g_{ij}$, and after some algebraic manipulations we get the following condition:

$$0 \leq \tau \left(h_{ij} + \frac{2g_{ij}}{h^2} (2-\gamma) \right) \leq 2.$$

Since γ is the sum of two cosines, then $0 \leq (2-\gamma) \leq 4$. This leads to

$$0 \leq \tau \left(h_{ij} + \frac{2g_{ij}}{h^2} (2-\gamma) \right) \leq \tau \left(h_{ij} + \frac{8g_{ij}}{h^2} \right) \leq 2.$$

The stability condition for the explicit scheme is then given by:

$$\tau \leq \frac{2h^2}{\max_{ij}(h^2 h_{ij} + 8g_{ij})}. \quad (A.2)$$

- (b) *Semi-implicit scheme*: Using the same procedure, as in the explicit case, the sufficient stability condition for the difference scheme in (11) implies:

$$\left| \frac{1 - \tau h_{ij}}{1 + 2\tau r_{ij}(2-\gamma)} \right| \leq 1, \quad (A.3)$$

where γ is as above. Knowing that $(2-\gamma) \geq 0$ and $r_{ij} \geq 0$, it is easy to verify that the denominator is always ≥ 1 . Given $h_{ij} \geq 0$, the above stability condition finally leads to:

$$\tau \leq \frac{2}{h_{\max}}, \quad \text{where } h_{\max} = \max_{ij} h_{ij}. \quad (A.4)$$

- (c) *Implicit scheme*: For the implicit scheme given in (10), the von Neumann stability criterion gives us the following expression of ξ :

$$\xi = \frac{1}{1 + h_{ij}\tau + 2\tau r_{ij}(2-\gamma)}, \quad (A.5)$$

where γ is as above. The stability condition $|\xi| \leq 1$ reads:

$$|1 + h_{ij}\tau + 2\tau r_{ij}(2-\gamma)| \geq 1,$$

which is always satisfied as $(2-\gamma) \geq 0$ and both r_{ij} and $h_{ij} \geq 0$. Hence the implicit scheme is unconditionally stable.

References

- [1] M. Kass, A. Witkin, D. Terzopoulos, Snake: Active contour model, Int. J. Comput. Vision 1 (1987) 321–331.
- [2] L. Cohen, I. Cohen, Finite element methods for active contour models and balloons for 2D and 3D images, IEEE Trans. Pattern Anal. Mach. Intell. 15 (11) (1993) 1131–1147.
- [3] L. Cohen, On active contour models and balloons, CVGIP: Image Understanding 53 (1991) 211–218.
- [4] B. Leroy, I. Herlin, L. Cohen, Multi-resolution algorithm for active contour models, Twelfth International Conference on Analysis and Optimization of Systems, vol. 15, no. 11, 1996, pp. 58–65.
- [5] A.C. Jalba, M.H.F. Wilkinson, J.B.T.M. Roerdink, CPM: A deformable model for shape recovery and segmentation based on charged particles, IEEE Trans. Pattern Anal. Mach. Intell. 26 (10) (2004) 1–16.
- [6] R. Yang, M. Mirmehdi, X. Xie, A charged active contour based on electrostatics, Advanced Concepts for Intelligent Vision Systems, Lecture Notes on Computer Science, vol. 4179, Springer-Verlag, 2006, pp. 173–184 Sep..
- [7] L. Bing, S. Acton, Active contour external force using vector field convolution for image segmentation, IEEE Trans. Image Process. 16 (8) (2007) 2096–2106.

- [8] C. Xu, J. Prince, Snakes, shapes, and gradient vector flow, *IEEE Trans. Image Process.* 7 (3) (1998) 359–369.
- [9] C. Xu, J. Prince, Generalized gradient vector flow external forces for active contours, *Signal Process.* 71 (1998) 131–139.
- [10] C. Xu, Deformable Models with Application to Human Cerebral Cortex Reconstruction from Magnetic Resonance Images," Ph.D. Thesis, Johns Hopkins University, Baltimore, MD 21218, USA, 1999.
- [11] N. Paragios, A variational approach for the segmentation of the left ventricle in cardiac image analysis, *Int. J. Comput. Vision* 50 (2002) 345–362.
- [12] N. Paragios, O. Mellina-Gottardo, V. Ramesh, Gradient vector flow fast geometric active contours, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (3) (2004) 402–407.
- [13] O. Colliot, O. Camara, I. Bloch, Integration of fuzzy spatial relations in deformable models—Application to brain MRI segmentation, *Pattern Recognition* 39 (8) (2006) 1401–1414.
- [14] G. Zhu, S. Zhang, Q. Zeng, C. Wang, Gradient vector flow active contours with prior directional information, *Pattern Recognition Lett.* 31 (9) (2010) 845–856.
- [15] N. Ray, S.T. Acton, Motion gradient vector flow: An external force for tracking rolling leukocytes with shape and size constrained active contours, *IEEE Med. Imag.* 23 (12) (2004) 1466–1478.
- [16] Y. Guo, C.-C. Lu, Multi-modality image registration using mutual information based on gradient vector flow, *International Conference on Pattern Recognition*, Los Alamitos, CA, USA, vol. 3, 2006, pp. 690–700.
- [17] M. Hassouna, A. Farag, On the extraction of curve skeletons using gradient vector flow, in: *International Conference on Computer Vision*, Brazil, October 14–20, 2007, pp. 1–8.
- [18] R. Donner, B. Micusik, G. Lings, L. Szumilas, P. Peloschek, K. Friedrich, H. Bischof, Object localization based on Markov random fields and symmetry interest points, in: *MICCAI'07*, Springer-Verlag, Berlin, Heidelberg, 2007, pp. 460–468.
- [19] J. Stöttinger, R. Donner, L. Szumilas, A. Hanbury, Evaluation of gradient vector flow for interest point detection, in: G. Bebis et al.(Ed.), *Advances in Visual Computing*, vol. 5358, Springer, Berlin, Heidelberg, 2008, pp. 338–348.
- [20] J. Cheng, S. Foo, Dynamic directional gradient vector flow for snakes, *IEEE Trans. Image Process.* 15 (6) (2006) 1563–1571.
- [21] J. Tang, S. Millington, S. Acton, J. Crandall, S. Harwitz, Ankle cartilage surface segmentation using directional gradient vector flow snakes, *International Conference on Image Processing*, vol. IV, 2004, pp. 2745–2748.
- [22] Y. Wang, W. Chen, T. Yu, Y. Zhang, Hessian based image structure adaptive gradient vector flow for parametric active contours, in: *International Conference on Image Processing*, 2010, pp. 649–652.
- [23] J. Ning, C. Wu, S. Liu, S. Yang, NGVF: An improved external force field for active contour model, *Pattern Recognition Lett.* 28 (1) (2007) 58–63.
- [24] L. Guillot, C. Le Guyader, Extrapolation of vector fields using the infinity Laplacian and with applications to image segmentation, in: X.-C. Tai et al.(Ed.), *SSVM series Lecture Notes on Computer Science*, vol. 5567, Springer, Berlin, Heidelberg, 2009, pp. 87–99.
- [25] C. Li, J. Liu, M.D. Fox, Segmentation of external force field for automatic initialization and splitting of snakes, *Pattern Recognition* 38 (11) (2005) 1947–1960.
- [26] C. Tauber, H. Batatia, A. Ayache, Une méthode d'initialisation automatique pour contour actif paramétrique, application aux images échographiques, *GRETSI*, vol. 1, Louvain-la-Neuve, Belgique, 2005, pp. 521–525.
- [27] P. Tauber, H. Batatia, A. Ayache, A robust active contour initialization and gradient vector flow for ultrasound image segmentation, in: Katsushi Ikeuchi (Ed.), *IAPR*, Tsukuba, Japan, May 2005.
- [28] E. Vidholm, P. Sundqvist, I. Nyström, Accelerating the computation of 3D gradient vector flow fields, *International Conference on Pattern Recognition*, Los Alamitos, Ca, USA, vol. 3, 2006, pp. 677–680.
- [29] X. Han, C. Xu, J. Prince, Fast numerical scheme for gradient vector flow computation using a multigrid method, *IET Image Process.* 1 (1) (2007) 48–55.
- [30] D. Boukerroui, Efficient numerical schemes for gradient vector flow, in: *IEEE Sixteenth International Conference on Image Processing*, Cairo, Egypt, 2009, pp. 4057–4060.
- [31] D. Boukerroui, Schémas numériques de résolution du flot de vecteur gradient, in: *Traitement et Analyse de l'Information: Méthodes et Applications*, Hammamet, Tunisie, May, 2009, pp. 523–532.
- [32] G. Aubert, P. Kornprobst, *Mathematical Problems in Image Processing*, November, Springer Verlag, 2001.
- [33] J. Weickert, B. Romeny, M. Viergever, Efficient reliable schemes for nonlinear diffusion filtering, *IEEE Trans. Image Process.* 7 (3) (1998) 398–410.
- [34] R. Tavakoli, P. Davami, 2D parallel and stable group explicit finite difference method for solution of diffusion equation, *Appl. Math. Comput.* 188 (2) (2007) 1184–1192.
- [35] M. Dehghan, Finite difference procedures for solving a problem arising in modeling and design of certain optoelectronic devices, *Math. Comput. Simul.* 71 (1) (2006) 16–30.
- [36] V. Saulév, *Integration of Equations of Parabolic Type by the Method of Nets*, Pergamon Press, New York, 1964.
- [37] D. Barash, T. Schlick, M. Israeli, R. Kimmel, Multiplicative operator splitting in nonlinear diffusion: from spatial splitting to multiple timesteps, *J. Math. Image Vision* 19 (1) (2003) 33–48.
- [38] J. Ramos, Iterative and non-iterative, full and approximate factorization methods for multidimensional reaction-diffusion equations, *Appl. Math. Comput.* 174 (2) (2006) 1586–1608.
- [39] J. Barron, D. Fleet, S. Beauchemin, Performance of optical flow techniques, *Int J. Comput. Vision* 12 (1) (1994) 43–77.
- [40] G. Papandreou, P. Maragos, Multigrid geometric active contour models, *IEEE Med. Imag.* 16 (1) (2007) 229–240.
- [41] P. Mrázek, M. Navara, Selection of optimal stopping time for nonlinear diffusion filtering, *Int J. Comput. Vision* 52 (2–3) (2003) 189–203.
- [42] C. Xu, J.L. Prince, Global optimality of gradient vector flow, in: *Proceedings of Thirty-fourth Annual Conference on Information Sciences and Systems (CISS'00)*, 2000, pp. 1–2.
- [43] Y. Wang, J. Liang, Y. Jia, On the critical point of gradient vector flow snake, *Asian Conference on Computer Vision*, vol. II, 2007, pp. 754–763.

Djamal Boukerroui was born in Bejaïa (Algeria) in 1972. He received the B.S. degree in 1990, and the M.S. degree in electronics in July 1995 from ENP of Algiers (Algeria). He received the Ph.D. degree in Image Processing at CREATIS Laboratory of INSA of Lyon (France) in 2000. From March 2000 to August 2002 he was a Research Assistant at the Medical Vision Laboratory, Department of Engineering Science, University of Oxford where he worked on the analysis of echocardiographic image sequences. Since September 2002, he joined the Department of Information Processing Engineering of Compiègne University of Technology as Maître de Conférences, and he is a member of HEUDIASYC [research unit associated with CNRS (#6599)]. His current main research interests are in low level image processing and specifically its application in ultrasound image analysis.